

The Error-Pattern-Correcting Turbo Equalizer

Hakim Alhussien, *Member, IEEE*, and Jaekyun Moon, *Fellow, IEEE*

Abstract

The error-pattern correcting code (EPCC) is incorporated in the design of a turbo equalizer (TE) with aim to correct dominant error events of the inter-symbol interference (ISI) channel at the output of its matching Viterbi detector. By targeting the low Hamming-weight interleaved errors of the outer convolutional code, which are responsible for low Euclidean-weight errors in the Viterbi trellis, the turbo equalizer with an error-pattern correcting code (TE-EPCC) exhibits a much lower bit-error rate (BER) floor compared to the conventional non-precoded TE, especially for high rate applications. A maximum-likelihood upper bound is developed on the BER floor of the TE-EPCC for a generalized two-tap ISI channel, in order to study TE-EPCC's signal-to-noise ratio (SNR) gain for various channel conditions and design parameters. In addition, the SNR gain of the TE-EPCC relative to an existing precoded TE is compared to demonstrate the present TE's superiority for short interleaver lengths and high coding rates.

Index Terms

Inter-symbol interference, turbo equalization, dominant error events, error pattern correcting code, maximum-likelihood bit error rate bound, error weight enumerator, list decoding, decode channel.

This work has been submitted to the special issue of the IEEE Transactions on Information Theory titled: "Facets of Coding Theory: from Algorithms to Networks". This work was supported in part by the NSF Theoretical Foundation Grant 0728676. Hakim Alhussien was with the department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, Minnesota 55455, U.S.A. He is now with Link-A-Media Devices, Santa Clara, CA 95051, USA (e-mail:hakima@link-a-media.com). Jaekyun Moon was with the department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, Minnesota 55455, U.S.A. He is now with the department of Electrical Engineering at KAIST, Yuseong-gu, Daejeon, 305-701, Republic of Korea (e-mail:jaemoon@ee.kaist.ac.kr).

I. INTRODUCTION

The turbo code of [1], [2] has been utilized as a practical means to approach the inter-symbol-interference (ISI) channel capacity in what has been termed turbo equalization [3], [4], in which two recursive systematic convolutional codes (RSCCs) concatenated in parallel are concatenated serially to the ISI channel. Since then, the turbo equalization terminology has grown to encompass any soft-decodable code that is iteratively decoded by exchanging soft information with a channel matched detector. The family of turbo equalizers now includes low-density-parity check (LDPC) codes and turbo product codes (TPC). A standard turbo code is a parallel concatenation of convolutional codes (PCCC) connected by an interleaver, for which the probability of generating low Euclidean weight error events is considerably reduced by the action of the uniform interleaver. This in effect improves the overall system bit-error-rate (BER) in the low-to-medium signal-to-noise ratio (SNR) region. A PCCC is decoded by an iterative exchange of soft information between maximum *a posteriori* probability (MAP) decoders matched to the constituent RSCC decoders [5]. A turbo equalizer (TE) based on an iterative receiver composed of a PCCC soft decoder and a channel detector was discussed in [6]. A simpler serial concatenation of a single RSCC and a precoder through an interleaver was found to perform just as well in [7] for wireless communication applications, and in [8] and [9] for magnetic recording applications. Precoding makes the ISI channel appear recursive to the outer interleaved RSCC, where the non-precoded ISI channel can be viewed as an inner nonrecursive rate-1 convolutional code [4]. In this manner, precoding is essential to achieve better turbo gain in the low SNR region, i.e. “waterfall region”. This was first shown in the context of serially concatenated convolutional codes (SCCCs) in additive white Gaussian noise (AWGN) in [10], where it was demonstrated that the inner constituent convolutional code has to be recursive to achieve a turbo gain. Briefly afterwards, this was demonstrated for a SCCC-TE running on the dicode channel in [11]. The concatenation of precoding and RSCC through an interleaver works by enhancing the error weight “spectral thinning” effect, by which the frequency of low Euclidean distance errors is uniformly reduced.

We propose an alternate error-weight spectral shaping approach that aggressively targets the low end of the error Euclidean distance distribution, enhancing BER performance in the “error floor” region, while maintaining the waterfall region gain of conventional TE. The proposed

method is based on directly targeting the dominant error patterns of the channel, which are also the lowest Euclidean distance errors, via a matching error correction code, termed the error-pattern correction code (EPCC). The EPCC was first proposed to handle single dominant error event occurrence in [12] [13], and later enhanced in [14] and [15] to handle multiple error event occurrences. A practical EPCC-based turbo equalizer tailored to the magnetic recording application was first proposed in [16]. In our TE setup, the EPCC is matched to the non-precoded ISI channel and serves as an inner code for an outer interleaved RSCC. Since the EPCC maintains a substantial error correction power while having a high code rate that is close to 1, the hope is that the redistribution of redundancy between EPCC and the outer RSCC would improve overall system performance. In TE-EPCC decoding, the EPCC MAP decoder works iteratively with the outer RSCC MAP decoder to correct low Euclidean distance errors at the output of the channel's detector. This is compared to using a rate-1 precoder in the encoder side that prevents these errors from occurring in high frequency but can not eliminate them entirely.

In this work, we conduct an error-event weight analysis of EPCC enhanced TE to be able to predict an upper bound on the BER performance, and hence establish the advantage of incorporating EPCC in the error floor region. The derived upper bound on BER is for the maximum likelihood (ML) decoder of the concatenated system, which the practical decoder is assumed to approach at high SNRs. A few points are worth mentioning regarding the derivation of such a bound. First, the bound is based on the notion of a uniform interleaver, which essentially averages out the effect of good and bad instantaneous interleavers on the bound. The implication of this assumption on the analytic BER bound is that the particular choice of the practical interleaver is not a factor in our turbo system comparison herein. Second, the derivation of the bound presumes a maximum-likelihood decoder, which fails short of accurately describing the iterative turbo gain that is more pronounced at lower to medium SNR, where the analysis of turbo code performance at this lower SNR region remains largely an open problem. Incidentally, our proposed approach here based on probabilistic correction of low Euclidean distance errors is designed to work in the floor region where the bound is accurate. Finally, the bound assumes that coded data is i.i.d., which becomes a more realistic approximation as the code rate of the RSCC approaches unity.

The paper is organized as follows; In Section II we review the main concepts of EPCC code construction and decoding based on its algebraic properties; we also present EPCC design

examples that we later use in the simulation of Section VI. In Section III we present the encoder and decoder components of the conventional precoded and non precoded TEs and of the TE-EPCC. Section IV analyzes the ML BER performance of the TE-EPCC and the conventional TE based on the overall error weight spectrum of the coded channel. Furthermore, this section discusses an efficient method to evaluate the BER bound based on multinomial theory, assuming a single EPCC codeword per interleave. In Section V we explain the gain of the TE-EPCC over the TE in terms of the improved interleaver gain exponents of lower Euclidean-weight errors. Section VI discusses a practical method to implement TE-EPCC decoding that approaches the ideal ML decoder analyzed in the preceding section. Finally, The numerical results in Section VII corroborate our claims in a variety of channel conditions for a combination of decoder design parameters.

II. REVIEW OF THE ERROR-PATTERN-CORRECTING CODE

The cyclic codes described in [13] are based on construction of a generator polynomial $g(x)$ that gives rise to distinct syndrome sets for all targeted dominant error patterns. It has been shown that such a $g(x)$ can be obtained from the irreducible factors making up the polynomial representations of the dominant error patterns. The code can be further improved by introducing another factor in $g(x)$, namely, a primitive polynomial that is not already a factor of $g(x)$ [14]. The results are an increased code rate, improved single-error-pattern correction accuracy (via reduced miss-correction probability), and capability to correct some important multiple-pattern events based on a increased number of distinct syndrome patterns.

We start by constructing a cyclic code targeting the set of l_{max} dominant error events

$$\{e_k^{(1)}(x), e_k^{(2)}(x), \dots, e_k^{(l_{max})}(x)\}$$

represented as polynomials on $GF(2)$ that can occur at any starting position k in the codeword of length l_T . A syndrome of error $e^{(i)}(x)$ at position k is defined as $s_k^{(i)}(x) = e_k^{(i)}(x) \bmod g(x)$, with $g(x)$ being the generator polynomial of the code and \bmod the polynomial modulus operation. A syndrome set \mathbf{S}_i for error type $e^{(i)}(x)$ contains elements corresponding to all cyclic shifts of polynomial $e^{(i)}(x)$; elements of \mathbf{S}_i are thus related by $s_{k+j}^{(i)} \equiv x^j s_k^{(i)} \bmod g(x)$.

For unambiguous decoding of $e^{(i)}(x)$ and $e^{(j)}(x)$, $\forall \{i, j\}$, we must have $\mathbf{S}_i \cap \mathbf{S}_j = \emptyset$. This design requirement constrains $g(x)$ to have distinct greatest common divisors with all $e^{(i)}(x)$.

However, even if this constraint is satisfied, an element in S_i can still map to more than one position, i.e., the period of the syndrome set- and period of $g(x)$ - can be less than l_{max} . Moreover, this constraint is only sufficient but not necessary. Also, as shown in [13], there may exist a lower degree $g(x)$ that can yield distinct syndrome sets for the targeted error polynomials, resulting in a higher rate EPCC. A search method to find this $g(x)$ is already discussed in detail in [13] and [15].

We now describe the construction and properties of the EPCC that will be deployed throughout the paper in the design of different turbo systems based on EPCCs. We target the dominant error events of a generalized two tap ISI channel of the form $1 - \alpha D$, $0 < \alpha \leq 1$, for which the dicode and PR1 channels are special cases. When α is close to 1, the dominant errors are: $+$, $+-$, $+ - +$, etc., which have the polynomial representations: $e^{(1)}(x) = 1$, $e^{(2)}(x) = 1 + x$, $e^{(3)}(x) = 1 + x + x^2$, etc., i.e. polynomials on $GF(2)$ for which all powers of x have nonzero coefficients.

For the purpose of designing EPCC codes for use in the TE-EPCC, the component EPCC code rate should be very high. To maintain high rate, the EPCC codeword has to be extended to a few hundred bits, without proportionally increasing the number of parity bits required to achieve accurate single-error occurrence correction capability. Example EPCC codes are shown next, and the syndrome set periods of these codes are shown in Table I.

- (630,616) *EPCC*: Targeting error polynomials up to degree 9, we get the generator polynomial $g(x) = 1 + x^3 + x^5 + x^8$ of period 30, via the search procedure in [13]. Choosing a codeword length of 30, 10 distinct, non-overlapping syndrome sets are utilized to distinguish the 10 target errors. However, the resulting (30,22) EPCC has rate 0.73 which incurs high rate penalty. By multiplying the base EPCC generator polynomial by the primitive polynomial $1 + x + x^6$, which is not a factor of any of the targeted errors, we obtain the extended generator polynomial $g_e(x) = 1 + x + x^3 + x^4 + x^5 + x^8 + x^{11} + x^{14}$, which corresponds to the extended (630,616) EPCC code of rate 0.98, and 14 parity bits. Then, as shown in [13], syndrome sets S_1 , S_3 , S_7 , and S_9 have period 630 and thus can be decoded without ambiguity. On the other hand, syndrome sets S_2 , S_4 , S_6 , and S_8 have period 315, decoding to one of two positions. The worst would be S_5 of period 126, and S_{10} of period 63, which decode to 5 and 10 possible positions, respectively. Still, the algebraic decoder can quickly shrink the number of possible error positions to few positions by checking the

data support, and then would choose the one position with highest local reliability.

- *Shortened* (126, 112) EPCC: Shorter lower-rate EPCC codes can be obtained by shortening the (630, 616) EPCC. For example, a (126, 112) EPCC of rate 0.89 can be derived this way with all syndromes sets, excluding syndrome set S_{10} , having period 126, and thus are decodable without ambiguity.
- (210, 199) EPCC: To obtain short EPCC codes without jeopardizing the code rate through code shortening, we can target fewer error patterns in the code design. Targeting error polynomials up to degree 9, but excluding $e^{(7)}(x)$, we can extend the base generator polynomial $g(x) = 1 + x^3 + x^5 + x^8$ through its multiplication by the primitive polynomial $1 + x + x^3$, which we could not use before because its a factor of the polynomial representation of $e^{(7)}(x)$. The resulting code is a (210, 199) EPCC of rate 0.95, 11 parity bits, and extended generator polynomial $g_e(x) = 1 + x + x^4 + x^5 + x^9 + x^{11}$.

TABLE I

SYNDROME SET PERIODS OF VARIOUS EPCC CODES.

Target error	(630, 616) EPCC	(126, 112) EPCC	(210, 199) EPCC
1	630	126	210
$(1 + x)$	315	126	105
$(1 + x + x^2)$	630	126	70
$(1 + x)^3$	315	126	105
$(1 + x + x^2 + x^3 + x^4)$	126	126	42
$(1 + x)(1 + x + x^2)^2$	315	126	35
$(1 + x + x^3)(1 + x^2 + x^3)$	630	126	—
$(1 + x)^7$	315	126	105
$(1 + x + x^2)(1 + x^3 + x^6)$	630	126	70
$(1 + x)(1 + x + x^2 + x^3 + x^4)^2$	63	63	21

III. A TE INCORPORATING THE EPCC SISO DECODER

The structure of the conventional SCCC-TE is shown in Fig. 1(i). In the encoder side, a simple RSCC encodes the data stream, which is interleaved before being passed to the channel.

The concatenation of the convolutional code and ISI channel can be viewed in the context of turbo coding as a serial concatenation of an outer recursive code and an inner rate-1 non-recursive code through an interleaver. A polynomial-time iterative-type decoder can be designed based on the separation of the ML decoders of the inner and outer codes. The ML decoders iteratively exchange reliability information converging to the combined ML solution at high SNR. Separate ML detection and decoding can be realized via the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [17], the soft output Viterbi algorithm (SOVA) [18], or the minimum mean squared error (MMSE) soft-in soft-out (SISO) detector [19], [20]. The BER gain in TE, however, is most notable at low SNRs, and declines rapidly as SNR increases resulting eventually in the error floor phenomenon. The gain at low SNR is further enhanced by including a rate-1 recursive component in the path of the coded interleaved bit stream. This is shown in Fig. 1(ii), where the trellis of SOVA is now matched to the recursive rate-1 coded channel $\frac{1-\alpha D}{1\oplus D}$. By the action of the ideal uniform interleaver, the fraction of errors in the Hamming-weight error distribution of the RSCC resulting in low Euclidean weight errors in the channel trellis is greatly reduced. This, as a result, improves the BER at low to medium SNR, where the contribution of the profile of error Euclidean weights to the BER far exceeds the single contribution of the minimum of these weights.

A markedly different approach is proposed in the structure of Fig. 1(iii). The new method is based on replacing the rate-1 precoder with a high rate ECC that is designed to correct low Hamming weight errors that generate low Euclidean-weight trellis errors rather than constraining their incidence. Since the targeted errors possess low Hamming weights by design, this reduces the added complexity of encoding and decoding the EPCC, while the intrinsic channel property of these errors generating low Euclidean weight errors, particularly $d_E^2 = 2$, lowers the error floor at medium to high SNRs substantially. Nevertheless, since the practical decoder of the EPCC incurs some miscorrection, this new approach resembles a probabilistic “best effort” enhancement of $d_{E,min}^2$ that is achieved by correcting a sizable fraction of the originating Hamming weight errors. A soft-in soft-out (SISO) decoder of the EPCC is assumed in the iterative turbo loop. Since the EPCC is matched to the ISI channel, no interleaving should be present between the EPCC and the channel. On the other hand, an interleaver is essential between the EPCC and the outer RSCC.

While the $\frac{1}{1\oplus D}$ precoder and the $1 - \alpha D$ channel can be jointly decoded with no added

complexity by matching the trellis to the combined coded channel $\frac{1-\alpha D}{1+\oplus D}$, its impractical to realize a similar joint ML decoder of the channel and the EPCC. Hence, in Fig. 1(iii) separate SISO decoders of the channel and the EPCC are implemented.

IV. ERROR-RATE ANALYSIS OF TE-EPCC

In bounding the BER of the TE-EPCC, many of the basic steps and assumptions taken in [11] and [21] for bounding the BER of the conventional TE are utilized. To more closely reflect the practical recording channel, however, we apply our BER analysis to a generalized two-tap channel of the form $1 \pm \alpha D$. The dicode $(1 - D)$ and PR1 $(1 + D)$ channels are special cases corresponding to $\alpha = 1$. In the proposed approach, we show how the BER is function of the error Euclidean distance distribution of the overall system. Then, we argue for TE-EPCC's enhanced performance by the virtue of its reduction of occurrence frequencies of low Euclidean distances in the overall distance distribution; it will also be shown that these low Euclidean distance components of the distribution dominate the system BER. Following the notations of [11], the maximum likelihood (ML) union bound on word error rate of a block code of codebook size M , of equally likely codewords and AWGN of zero mean and variance σ^2 is

$$P_W \leq \frac{1}{M} \sum_{m=1}^M \sum_{\hat{m} \neq m} Q \left(\frac{\| \mathbf{x}_m - \mathbf{x}_{\hat{m}} \|}{\sigma} \right) \quad (1)$$

where m and \hat{m} are codewords separated by the Euclidean distance $\| \mathbf{x}_m - \mathbf{x}_{\hat{m}} \|$, and \mathbf{x}_m is the noiseless channel output for m . If there are T_{m,d_E} different codewords for which the corresponding noiseless channel outputs are at distance d_E from \mathbf{x}_m , then we can write (1) as:

$$\begin{aligned} P_W &\leq \frac{1}{M} \sum_{m=1}^M \sum_{d_E=1}^{\infty} T_{m,d_E} Q \left(\frac{d_E}{\sigma} \right) \\ &= \sum_{d_E=d_E^{min}}^{\infty} \bar{T}(d_E) Q \left(\frac{d_E}{\sigma} \right) \end{aligned} \quad (2)$$

where $\bar{T}(d_E)$ is the average number of codewords at Euclidean distance d_E from a given codeword, with the distance measured at the channel output. The associated BER can be shown to be

$$P_b \leq \sum_{d_E=d_E^{min}}^{\infty} \frac{\bar{T}(d_E) \bar{w}(d_E)}{K} Q \left(\frac{d_E}{\sigma} \right) \quad (3)$$

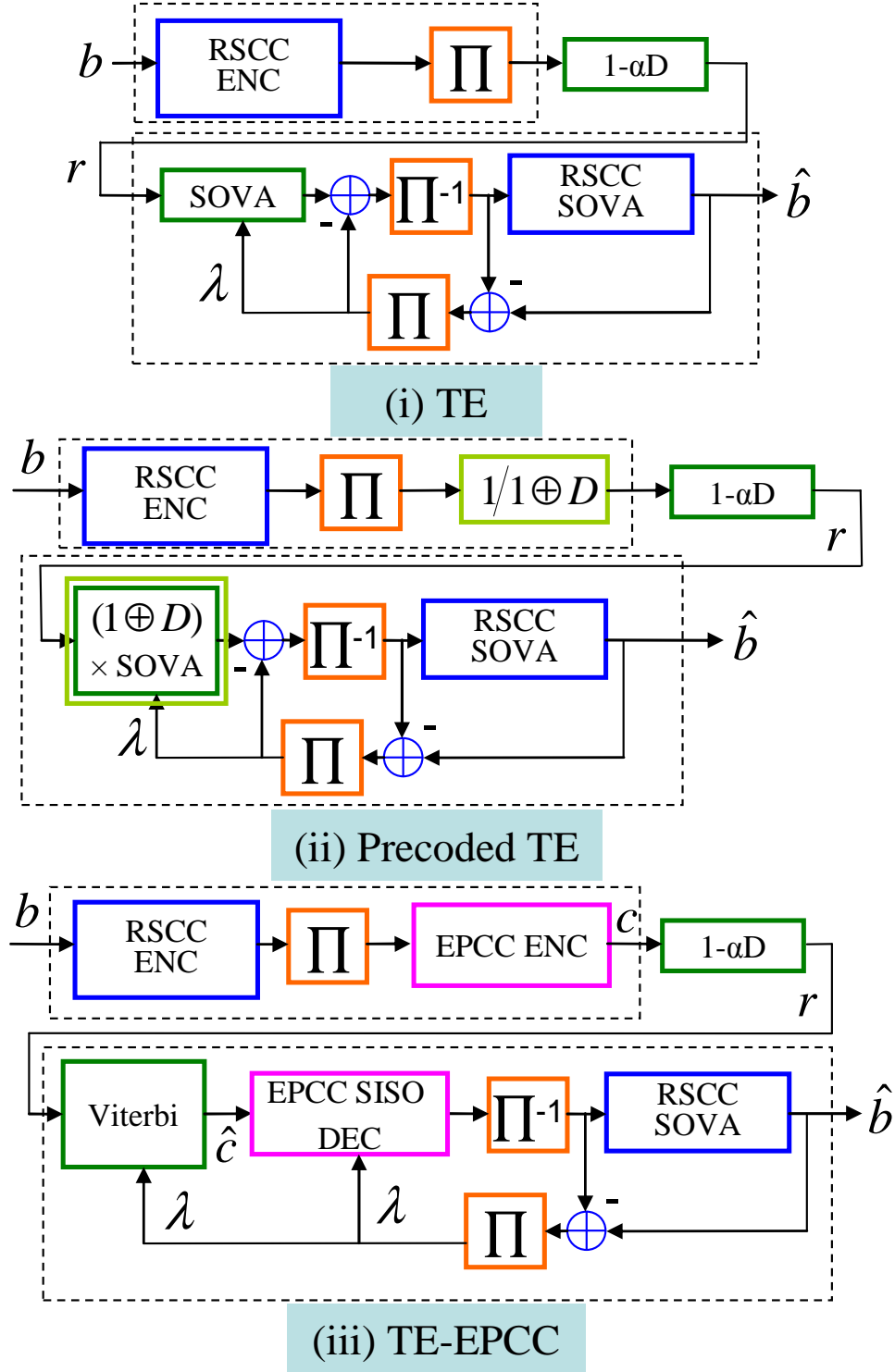


Fig. 1. Block diagrams: (i) TE, (ii) precoded TE, (iii) TE-EPCC.

where K is the number of information bits per codeword sequence and $\overline{w}(d_E)$ is the average Hamming distance from a given information word to competing information words located at d_E away, with the Euclidean distance measure based on noiseless channel outputs of the corresponding codewords. We next show how $\overline{T}(d_E)$ is related to the outer code Hamming weight enumerator $A^o(d)$ and the error event characteristics of the channel.

A. Error Event Analysis of the $1 - \alpha D$ Channel

A trellis section of the $1 - \alpha D$ channel with no precoding is shown in Fig. 2. The branch label c_i/x_i signifies the coded input bit to the channel, and the corresponding channel output, respectively. Following the same notation as in [11], any error word \mathbf{f} with Hamming weight $d = d^H(\mathbf{f})$ can be uniquely decomposed into a concatenation of disjoint error patterns \mathbf{f}_j , $j = 1, \dots, m$, where the index j signifies the order of occurrence of the error pattern of Hamming weight d_j^H in the codeword. Error patterns \mathbf{f}_j , $j < m$, correspond to simple closed error events on the trellis that diverge from and remerge into the correct path without sharing any of the states in between. However, two scenarios can occur when $j = m$: either \mathbf{f}_m remerges with the correct path (closed \mathbf{f}_m) or the boundary of the codeword is reached while the two paths are still diverged (open \mathbf{f}_m).

In the $1 - \alpha D$ channel trellis, diverging branches result in a Euclidean distance separation of 1 each, while remerging branches result in a squared Euclidean distance separation of α^2 each. Moreover, crossing branches accumulate a squared distance separation of $(1 + \alpha)^2$, while parallel branches accumulate a separation of $(1 - \alpha)^2$. This means that parallel branches result in a lower Euclidean distance separation compared to crossing branches in the Euclidean distance distribution when $0 < \alpha \leq 1$.

Hence, two error pattern classes are distinguishable according to their accumulate Euclidean distance. The first class, shown in Fig. 2b, has a squared distance $d_E^2(\mathbf{f}_j) = 1 + \mu\alpha^2 + (d_j^H - 1) \times (1 - \alpha)^2$ where d_j^H is the Hamming weight of the error event \mathbf{f}_j , and $\mu=0$ or 1 depending on the event being open or closed, respectively. This class of error patterns is denoted by χ^{dom} and is called “the dominant error class”, for which all branches, other than the diverging and remerging branches, are parallel. The dominant error class accounts for most of the channel bit errors due to the low Euclidean distance between the correct and erroneous paths. On the other hand, the second class, shown in Fig. 2c, has both parallel and crossing branching, and hence its members

have Euclidian distance $d_E^2(\mathbf{f}_j) = 1 + \mu\alpha^2 + \lambda_{cr} \times (1 + \alpha)^2 + (d_j^H - 1 - \lambda_{cr}) \times (1 - \alpha)^2$, where λ_{cr} is the number of crossing branches. The second class contributes much less to the overall system BER, and thus we call it “the non-dominant error class”, which is denoted by $\tilde{\chi}^{dom}$. By the same line of argument, the same two classes are distinguishable for the PR1 channel, which is a special case of $1 + \alpha D$ at $\alpha = 1$. The only difference is that error events with all crossing branches now generate the class χ^{dom} .

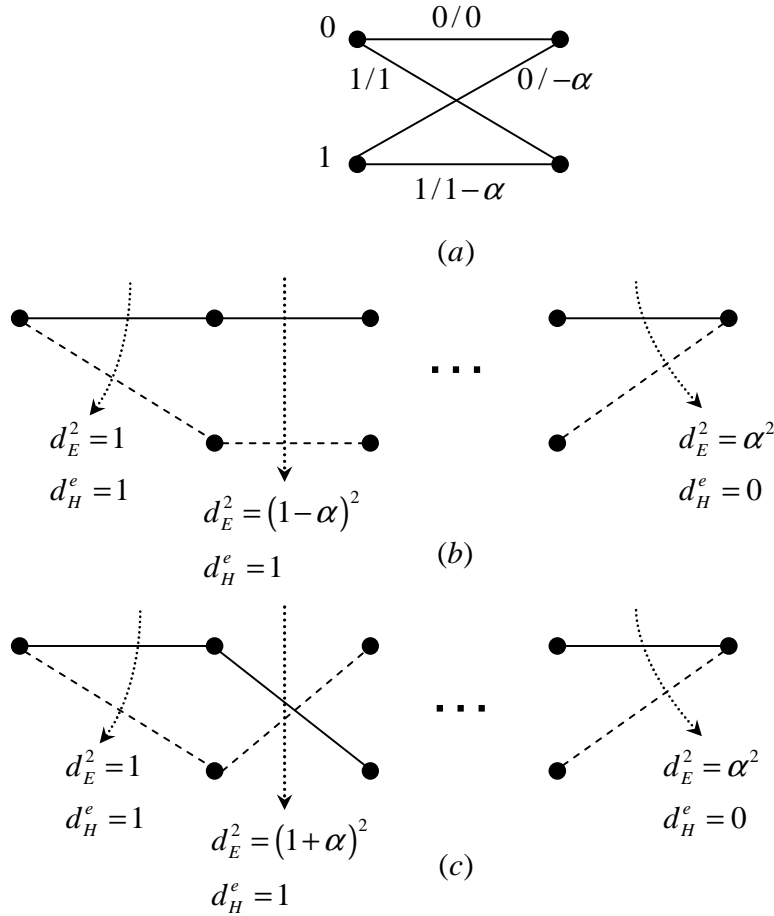


Fig. 2. (a) Trellis section for a non-precoded generalized two-tap ISI channel $(1 - \alpha D)$, (b) dominant error patterns, (c) non-dominant error patterns.

We design an error-pattern correcting code (EPCC) capable of correcting error codewords \mathbf{f} that are decomposable into disjoint error patterns \mathbf{f}_j that all belong to the dominant error class, i.e. $\mathbf{f}_j \in \chi^{dom}, \forall j$. In order to evaluate the BER performance of EPCC we need to find the new Euclidean distance distribution modified by EPCC. However, it would be easier to first find

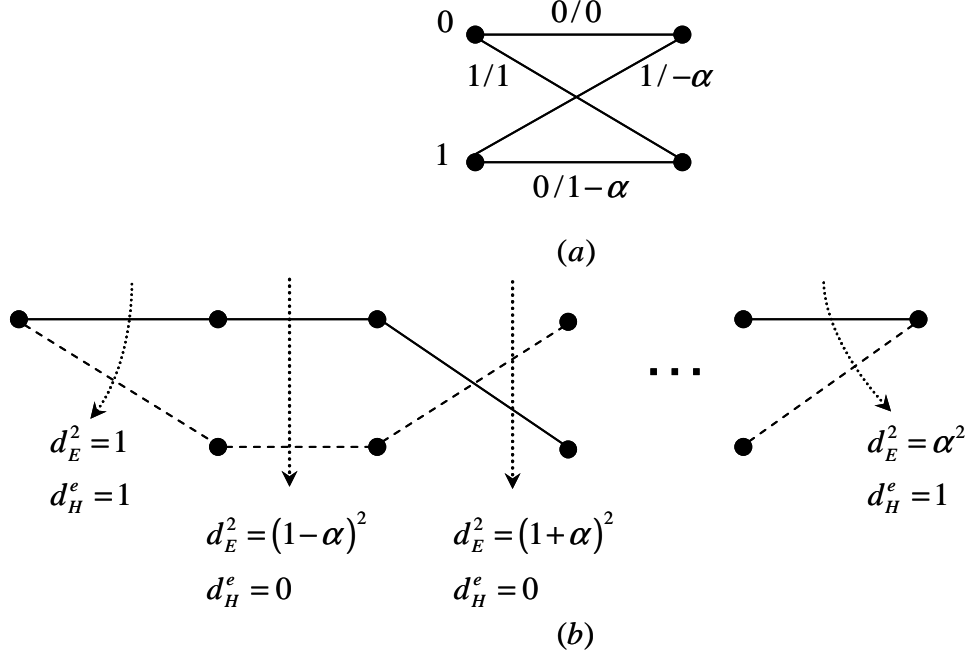


Fig. 3. (a) Trellis section for a $\frac{1}{1 \oplus D}$ precoded generalized two-tap ISI channel $(1 - \alpha D)$, (b) weight characterization of an error pattern

the Euclidean distance distribution before EPCC correction is turned on. We assume throughout that code bit values are i.i.d and equiprobable, which is a valid assumption for high rate codes. Suppose an error word \mathbf{f} , of Hamming weight $d^H(\mathbf{f}) = d$, is composed of m_{dom} error patterns $\mathbf{f}_j \in \chi^{dom}$, and $\tilde{m}_{dom} = m - m_{dom}$ error patterns $\mathbf{f}_j \in \tilde{\chi}^{dom}$. A dominant error pattern \mathbf{f}_j of length $l_j = d_H(\mathbf{f}_j)$ will have probability $(\frac{1}{2})^{l_j-1}$. On the other hand, a non-dominant error pattern \mathbf{f}_j of length l_j and λ_{cr} crossing branches will have a probability of $\binom{l_j-1}{\lambda_{cr}} (\frac{1}{2})^{l_j-1}$. Therefore, the probability distribution of $d_E^2(\mathbf{f})$ is given by:

$$\Pr(d_E|d, m) = \begin{cases} \binom{d-m}{\lambda_{cr}} \left(\frac{1}{2}\right)^{d-m}, & \lambda_{cr} > 0 \text{ integer,} \\ & m_{dom} < m. \\ \left(\frac{1}{2}\right)^{d-m_{dom}}, & \lambda_{cr} = 0, \\ & m_{dom} = m. \\ 0, & \text{otherwise.} \end{cases}$$

$$\lambda_{cr} = \frac{d_E^2 - (1-\alpha)^2 d - 2\alpha m + \mu\alpha^2}{4\alpha} \quad (4)$$

which is the conditional probability of an error word of Euclidean distance d_E^2 , given that its

Hamming weight is d , and has m multiple error pattern occurrences, of which m_{dom} belong to χ^{dom} .

If we examine the precoded $1 - \alpha D$ trellis in Fig. 3, we note that a nonzero Hamming error results in the diverging of a single error event that remerges only on the occurrence of another Hamming error, while all the in between error branches have zero Hamming weights, wether crossing or parallel. We also note that an even d^H compound Hamming error decomposes into $\frac{d^H}{2}$ closed single errors, while an odd d^H compound error decomposes into $\lfloor \frac{d^H}{2} \rfloor$ closed errors and a boundary error. Moreover, diverging and remerging branches have $d_E^2 = 1$ and $d_E^2 = \alpha^2$, respectively, while parallel and crossing branches have $d_E^2 = (1 - \alpha)^2$ and $d_E^2 = (1 + \alpha)^2$, respectively. This means that, by invoking the random uniform interleaver assumption, the probability of a single long error event of $d^H = \{1, 2\}$ producing a low Euclidean weight error declines rapidly as the interleaver length is increased, since the probability of an all parallel error event declines accordingly. The Euclidean distance of a multiple error event of Hamming weight d , λ_{cr} crossing branches, and total length L is:

$$d_E^2 = \lceil \frac{d}{2} \rceil + \lfloor \frac{d}{2} \rfloor \alpha^2 + (1 + \alpha)^2 \lambda_{cr} + (1 - \alpha)^2 (L - d - \lambda_{cr})$$

Therefor

$$\begin{aligned} \Pr(d_E | d, L) &= \begin{cases} \binom{L-d}{\lambda_{cr}} \left(\frac{1}{2}\right)^{L-d}, & \lambda_{cr} > 0 \text{ integer.} \\ 0, & \text{otherwise.} \end{cases} \\ \lambda_{cr} &= \frac{d_E^2 - \lceil \frac{d}{2} \rceil - \lfloor \frac{d}{2} \rfloor \alpha^2 - (1 - \alpha)^2 (L - d)}{4\alpha} \end{aligned} \quad (5)$$

B. Error Euclidean Distance Distribution of TE-EPCC

We now develop a method to construct the error Euclidean distance distribution of TE-EPCC, for which the comparable distance distribution of TE is a special case where EPCC is turned off. Consider a serial concatenation of an EPCC and an interleaved recursive systematic convolutional code (RSCC) of length N . There are L_c EPCC subcodes in each interleave, each of length $N_c = \frac{N}{L_c}$, where an EPCC can correct up to m_c multiple occurrences per subcode provided that they all belong to the target set of correctable errors. The target set is $\{\mathbf{f}_j : \mathbf{f}_j \in \chi^{dom}, d^H(\mathbf{f}_j) \leq$

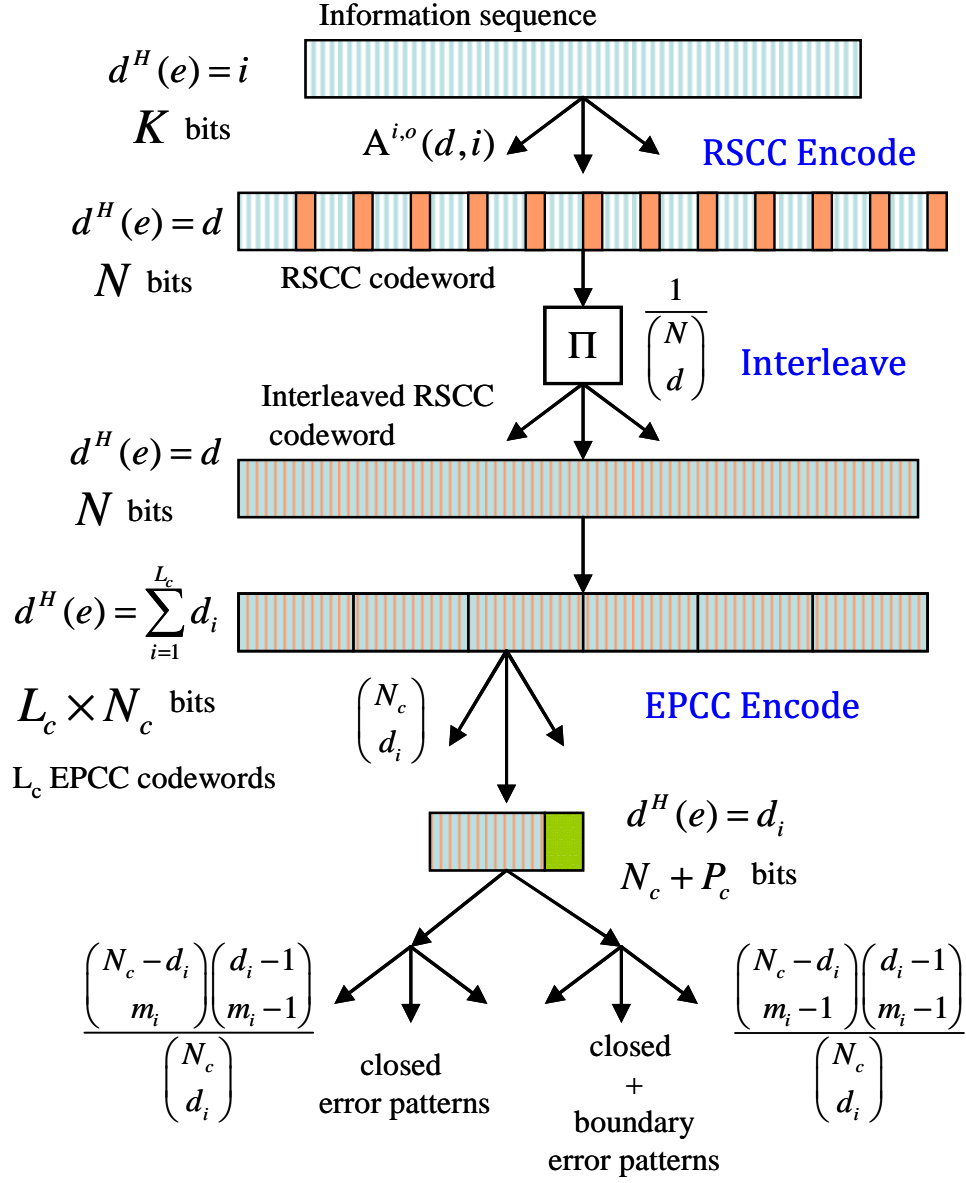


Fig. 4. Sketch of the method to derive $\bar{T}(d_E)$.

$d_c\}$, where d_c is the maximum length correctable error from χ^{dom} . An error in the RSCC codeword of hamming weight d is mapped by the uniform interleaver into all possible $\binom{N}{d}$ interleaved error words \mathbf{f} with equal probability. The interleaved error word divides into L_c EPCC subcodes, each receiving error word $\mathbf{f}^{(i)}$, $i = 1, \dots, L_c$, of Hamming weights d_1, d_2, \dots, d_{L_c} . Each EPCC error word $\mathbf{f}^{(i)}$ of Hamming weight d_i decomposes into m_i disjoint error pattern occurrences. In the previous section, we found the conditional probability $P(d_E|d, m)$ given

the error Hamming weight and number of multiple errors m for a single subcode interleave. To derive the Euclidean distance distribution for a codeword that is divisible into L_c subcodes, we are also required to evaluate the conditional probability of the decompositions m_i given the EPCC subcode hamming weights d_i . The conditional Euclidean distance probability distribution can be expanded as follows:

$$\begin{aligned}
\Pr(d_E|d) &= \Pr(d_E|d, d_1, \dots, d_{L_c}) \times \Pr(d_1, \dots, d_{L_c}|d) \\
&= \Pr(d_E|d, d_1, \dots, d_{L_c}, m, m_1, \dots, m_{L_c}) \\
&\quad \times \Pr(m, m_1, \dots, m_{L_c}|d_1, \dots, d_{L_c}, d) \\
&\quad \times \Pr(d_1, \dots, d_{L_c}|d).
\end{aligned} \tag{6}$$

Since errors in the L_c EPCC subcodes are disjoint, (6) becomes:

$$\begin{aligned}
\Pr(d_E|d) &= \sum_{d_1=0}^d \dots \sum_{\substack{d_{L_c}=0 \\ d=\sum_{i=1}^{L_c} d_i}}^d \Pr(d_1, \dots, d_{L_c}|d) \\
&\quad \sum_{m=1}^d \sum_{\substack{m_1=0 \\ m=\sum_{i=1}^{L_c} m_i}}^{d_1} \dots \sum_{m_{L_c}=0}^{d_{L_c}} \Pr(d_E|d, m) \prod_{i=1}^{L_c} \Pr(m_i|d_i)
\end{aligned} \tag{7}$$

The joint conditional probability $\Pr(d_1, \dots, d_{L_c}|d)$ in (7) is the probability of dividing the $\binom{N}{d}$ possible instants of the interleaved error word \mathbf{f} , of Hamming weight d , into the error word sequence $\mathbf{f}^{(i)}$ with associated Hamming weight sequence d_i , and is given by

$$\Pr(d_1, \dots, d_{L_c}|d) = \frac{\binom{N_c}{d_1} \times \binom{N_c}{d_2} \dots \times \binom{N_c}{d_{L_c}}}{\binom{N}{d}}. \tag{8}$$

Given that there are d_i errors in EPCC subcode i , there exists $\binom{d_i-1}{m_i-1}$ ways by which the Hamming weight d_i error is decomposed into m_i multiple error pattern occurrences, each of length at least 1. Of these m_i occurrences, $\mathbf{f}_{m_i}^{(i)}$ can be either open or closed. An open error event in this context lies on the boundary of the EPCC subcode's data and parity fields. A boundary error event is defined this way since we do not count error patterns in the EPCC codeword's parity towards the total BER, where this field is discarded before passing the decoded data to the outer interleaved RSCC decoder. By examining the trellis we note that boundary errors contribute a

squared Euclidean distance separation that is α^2 less than identical length closed errors that are totally encapsulated by the current subcode i data field. Furthermore, there are only $\binom{N_c-d_i}{m_i-1}$ ways by which the disjoint m_i error patterns of error word $\mathbf{f}^{(i)}$ can be arranged in the current subcode i , given the subcode has a boundary error. Two disjoint error occurrences in the trellis are separated at least by the error free distance of the channel, which equals 1 for $1 \mp \alpha D$ ISI channels. The number of possible arrangements of m_i errors is computed given the fact that the last error pattern occurs at the boundary. Assuming errors can occur on and off the boundary, the total number of possible error pattern arrangements becomes $\binom{N_c-d_i+1}{m_i}$. Given that the EPCC parity field is long enough, boundary errors have very low probability of spanning the data fields of adjacent EPCC subcodes, and hence, such events are independent among different subcodes.

So, given μ_i , there are $\binom{N_c-d_i}{m_i-\mu_i}$ ways by which the m_i error patterns, composing $\mathbf{f}^{(i)}$, can be arranged in a subcode i , and since there are $\binom{N_c}{d_i}$ possible error words $\mathbf{f}^{(i)}$, we get

$$\Pr(m_i|\mu_i, d_i) = \frac{\binom{N_c-d_i}{m_i-\mu_i} \times \binom{d_i-1}{m_i-1}}{\binom{N_c}{d_i}}. \quad (9)$$

A pictorial depiction of the derivation method explained above is shown in Figure 4. Substituting (4), (9), and (8) into (7), we get an expression for the distribution of error Euclidean distances while EPCC is turned off as:

$$\begin{aligned} \Pr(d_E|d) &= \sum_{m=1}^d \sum_{\mu=1}^{L_c} \\ & d = \sum_{i=1}^{L_c} d_i, m = \sum_{i=1}^{L_c} m_i, \mu = \sum_{i=1}^{L_c} \mu_i \\ & m, d, \mu, \alpha : \frac{d_E^2 - 2\alpha m + \mu\alpha^2 - (1-\alpha)^2 d}{\alpha} = 0 \bmod 4 \\ & \frac{1}{\binom{N}{d}} \prod_{i=1}^{L_c} \sum_{d_i=0}^d \sum_{m_i=0}^{d_i} \sum_{\mu_i=0}^1 \left(\frac{d-m}{\frac{d_E^2 - 2\alpha m + \mu\alpha^2 - (1-\alpha)^2 d}{4\alpha}} \right) \\ & \left(\frac{1}{2} \right)^{d-m} \binom{N_c - d_i}{m_i - \mu_i} \binom{d_i - 1}{m_i - 1} \end{aligned} \quad (10)$$

where we define $\binom{0}{0} = 1$. In addition, the Euclidean distance distribution can be decomposed into two components: a component $\Pr(d_E|d, \mathcal{C})$ associated with error words that are correctable by the L_c EPCC subcodes, and the complimentary component $\Pr(d_E|d, \tilde{\mathcal{C}})$ associated with non-correctable error words. In this case, the Euclidean distance probability distribution of non-

correctable error words escaping TE-EPCC is given by

$$\Pr(d_E|d, \tilde{\mathcal{C}}) = \Pr(d_E|d) - \Pr(d_E|d, \mathcal{C}) \quad (11)$$

while the correctable component is given by:

$$\begin{aligned} \Pr(d_E|d, \mathcal{C}) &= \sum_{m=1}^d \sum_{\mu=1}^{L_c} \\ &d = \sum_{i=1}^{L_c} d_i, m = \sum_{i=1}^{L_c} m_i, \mu = \sum_{i=1}^{L_c} \mu_i \\ &m, d, \mu, \alpha : d_E^2 = 2\alpha m - \mu\alpha^2 + (1 - \alpha)^2 d \\ &\frac{1}{\binom{N}{d}} \prod_{i=1}^{L_c} \sum_{d_i=0}^{\min(d, d_c)} \sum_{m_i=0}^{\min(d_i, m_c)} \sum_{\mu_i=0}^1 \\ &\left(\frac{1}{2}\right)^{d-m} \binom{N_c - d_i}{m_i - \mu_i} \binom{d_i - 1}{m_i - 1} \end{aligned} \quad (12)$$

where for the sake of simplicity, we assumed that an EPCC subcode i could correct an error word $\mathbf{f}^{(i)}$ if $d_H(\mathbf{f}^{(i)}) \leq d_c$, which is actually a worst case scenario that occurs only if $m_i = 1$. Although this assumption would result in a slightly pessimistic prediction of the EPCC correction power, it allows us to avoid a substantially more complicated derivation. To obtain the bound on the bit error probability, we need to express the error Euclidean distance enumerators as a function of the error Euclidean distance probability distribution given by (11). We note that the average Euclidean weight enumerator associated with the uncorrectable set of error words $\tilde{\mathcal{C}}$ is given by:

$$\overline{T}(d_E, \tilde{\mathcal{C}}) = \sum_{d=d_{\min}}^N \mathbf{A}^o(d) \Pr(d_E|d, \tilde{\mathcal{C}}) \quad (13)$$

while the average information input hamming distance to codewords at squared Euclidean distance d_E^2 is given by:

$$\overline{w}(d_E, \tilde{\mathcal{C}}) = \frac{1}{\overline{T}(d_E, \tilde{\mathcal{C}})} \sum_{d=d_{\min}}^N \mathbf{A}^o(d) \overline{\mathbf{A}}^i(d) \Pr(d_E|d, \tilde{\mathcal{C}}) \quad (14)$$

where $\mathbf{A}^o(d)$ represents the number of RSCC codeword sequences of weight d , and $\overline{\mathbf{A}}^i(d)$ represents the average input Hamming weight of RSCC codewords of weight d , and are related by

$$\overline{\mathbf{A}}^i(d) = \frac{\sum_i i \mathbf{A}^{i,o}(d, i)}{\mathbf{A}^o(d)}. \quad (15)$$

where $\mathbf{A}^{i,o}(d, i)$ is the number of codeword sequences of weight d that originated from weight i information sequences. Details on how to find these marginal error weight enumerators can be found in [22] for different puncturing rates and encoder connection polynomials. By substituting $\overline{T}(d_E, \tilde{\mathcal{C}})$, given by (13), and $\overline{w}(d_E, \tilde{\mathcal{C}})$, given by (14), in (3), we get an upper bound on the average BER of TE-EPCC as function of $\Pr(d_E|d, \tilde{\mathcal{C}})$:

$$P_b \leq \sum_{d_E=d_E^{min}}^{\infty} \sum_{d=d_{min}}^N \frac{\mathbf{A}^o(d) \overline{\mathbf{A}}^i(d) \Pr(d_E|d, \tilde{\mathcal{C}})}{K} Q\left(\frac{d_E}{\sigma}\right). \quad (16)$$

In Appendix A we show how these bounds simplify for the simple case when $L_c = 1$, i.e. employing one EPCC subcode per interleave. Also, we extend the BER bound derived in [11] for the precoded decode channel to the generalized case $1 - \alpha D$. Finally, by using an exponential-type approximation of the Q function, we show in appendix A that the BER bounds of the TE-EPCC, the non-precoded TE, and the precoded TE can be expressed as single infinite sums, with the Hamming weight of the RSCC error as the sum index.

C. Efficient Computation of the Euclidean Distance Enumerator for $L_c > 1$ EPCC

A more compact and efficient method is derived here to evaluate the multiple summations in equations (12) and (10), which are used to compute the BER bound in (16). We first define a probability enumerator for subcode i for all possible values of the parameters d_i , m_i and μ_i , which is given by the multinomial

$$\begin{aligned} \Lambda(D, M, \Upsilon; m_{max}, d_{max}) = & \\ 1 + \sum_{\mu_i=0}^1 \sum_{d_i=1}^{d_{max}} \sum_{m_i=1}^{\min(d_i, m_{max})} \left(\frac{1}{2}\right)^{d_i-m_i} & \\ \binom{N_c - d_i}{m_i - \mu_i} \binom{d_i - 1}{m_i - 1} D^{d_i} M^{m_i} \Upsilon^{\mu_i} & \end{aligned} \quad (17)$$

where the $D^0 M^0 \Upsilon^0 = 1$ monomial term corresponds to the case when there are no errors in the specified subcode, and $\mu_i = \{0, 1\}$ is the number of boundary errors per subcode. As a result, the probability enumerator for the entire interleave composed of L_c EPCC subcodes is given by

$$\Lambda^{L_c}(D, M, \Upsilon; m_{max}, d_{max})$$

given that only d_{max} -weight error words $\mathbf{f}^{(i)}$ composed of m_{max} disjoint error patterns can occur per EPCC subcode, where d_{max} and m_{max} are unbounded from above if EPCC correction is turned off. The advantage of this approach is that polynomial multiplication, or the more general multinomial multiplication, can be performed efficiently by symbolic manipulators, such as MapleTM, speeding up the evaluation of (12) and (10). Utilizing the compact, and easy-to-compute, probability enumerator, we can now express the bound on the bit error rate of the TE-EPCC as:

$$\begin{aligned}
P_b \leq & \frac{1}{K} \sum_{d_E=d_{min}}^{\infty} Q\left(\frac{d_E}{\sigma}\right) \sum_{d=d_{min}}^N \frac{\mathbf{A}^o(d)\overline{\mathbf{A}}^i(d)}{\binom{N}{d}} \sum_{\mu=0}^{L_c} \\
& \sum_{m=1}^d \binom{d-m}{\lambda_{cr}} [\Lambda^{L_c}(\mathbf{D}, \mathbf{M}, \Upsilon; \infty, \infty)]_{d,m,\mu} \\
& m: \lambda_{cr} \geq 0, \lambda_{cr} \in \mathbf{N} \\
& - \sum_{m=1}^d [\Lambda^{L_c}(\mathbf{D}, \mathbf{M}, \Upsilon; m_c, d_c)]_{d,m,\mu} \\
& m: \lambda_{cr}=0 \\
\lambda_{cr} = & \frac{d_E^2 - 2\alpha m + \mu\alpha^2 - (1-\alpha)^2 d}{4\alpha}
\end{aligned} \tag{18}$$

where the probability enumerator for a correctable EPCC codeword is approximated by

$$\Lambda^{L_c}(\mathbf{D}, \mathbf{M}, \Upsilon; m_c, d_c),$$

for an EPCC of maximum correction power m_c per subcode, and \mathbf{N} is the set of natural numbers.

V. INTERLEAVER GAIN EXPONENT OF TE-EPCC

To gain insight into how EPCC enhances TE performance, we pursue an analytic approach to study the mechanism by which EPCC reduces the multiplicity of low Euclidean distance errors. For this, we limit our investigation to the decode channel, for which the spectrum of the Euclidean distance is comprised only of integer values of d_E^2 , and hence there are a fewer values that d_E can take in the lower range of the spectrum. The error probability shown in (3) can be lowered by 1) increasing the minimum Euclidean distance between error words, a traditional approach, or 2) reducing the multiplicity of low Euclidean distance errors, as in the

TABLE II

INTERLEAVER GAIN EXPONENT OF THE CONVENTIONAL NON-PRECODED TE VS THE TE-EPCC, $d_E^2 = \{1, 2, 3, 4\}$.

$d_E^2 = 1$	Error pattern classes	TE	TE-EPCC
$m = 1$ $\mu = 1$ $d = 2 \rightarrow d_T$		N^{-2} (1)	N^{-11} $(\frac{155925}{4})$
$d_E^2 = 2$	Error pattern classes	TE	TE-EPCC
$m = 1$ $\mu = 0$ $d = 2 \rightarrow d_T$		N^{-1} (1)	N^{-10} $(\frac{155925}{4})$
$d_E^2 = 3$	Error pattern classes	TE	TE-EPCC
$m = 2$ $\mu = 1$ $d = 2 \rightarrow d_T$		N^{-1} (2)	N^{-10} (779625)
$d_E^2 = 4$	Error pattern classes	TE	TE-EPCC
$m = 2$ $\mu = 0$ $d = 2 \rightarrow d_T$		N^0 (1)	N^{-9} $(\frac{779625}{2})$

turbo coding paradigm pioneered by Berrou *et al.* [1]. In turbo coding, the coefficients of the error function for low Euclidean distances are an inverse function of the interleaver size, N . For this reason, turbo coding gain is often referred to as interleaver gain. At a more detailed level, for the frequency of low weight errors to asymptotically approach zero as the interleaver size tends to infinity, the exponent of the interleaver size in the corresponding error coefficients should be less than zero. Therefore, we can argue for the advantage of incorporating an EPCC in TE, by showing how it works to decrease the exponent of N^δ well below zero, especially for low Euclidean distance errors. We call δ the interleaver gain exponent. First, we isolate the exponent of N in the expression of BER for TE and TE-EPCC. The BER expression of the

TABLE III

INTERLEAVER GAIN EXPONENT OF THE CONVENTIONAL NON-PRECODED TE VS THE TE-EPCC, $d_E^2 = \{5, 6\}$.

$d_E^2 = 5$	Error pattern classes	TE	TE-EPCC
$m = 3$ $\mu = 1$ $d = 3 \rightarrow d_T$		N^{-1} (3)	N^{-9} (3508313)
$d_E^2 = 5$	Error pattern classes	TE	TE-EPCC
$m = 1$ $\mu = 1$ $d = 2 \rightarrow d_T$		N^{-2} (1)	N^{-2} (1)
$d_E^2 = 6$	Error pattern classes	TE	TE-EPCC
$m = 3$ $\mu = 0$ $d = 3 \rightarrow d_T$		N^0 (1)	N^{-8} (1169438)
$d_E^2 = 6$	Error pattern classes	TE	TE-EPCC
$m = 1$ $\mu = 0$ $d = 2 \rightarrow d_T$		N^{-1} (1)	N^{-1} (1)

conventional TE (EPCC turned off) for $\alpha = 1$ is:

$$\begin{aligned}
 P_b \leq & \frac{1}{K} \sum_{d_E=1}^{\infty} Q\left(\frac{d_E}{\sigma}\right) \sum_{d=2}^{d_T} \frac{\mathbf{A}^o(d) \overline{\mathbf{A}}^i(d)}{\binom{N}{d}} \\
 & \sum_{\mu=0}^1 \sum_{\substack{m=1 \\ m: d_E^2 - 2m + \mu = 0 \bmod 4}}^d \binom{d-m}{\frac{d_E^2 - 2m + \mu}{4}} \\
 & \left(\frac{1}{2}\right)^{d-m} \binom{N-d}{m-\mu} \binom{d-1}{m-1}
 \end{aligned} \tag{19}$$

where $d_T \ll N$ is the truncated maximum error weight. We truncated the Hamming error weight d since large weight errors correspond to larger Euclidean distances which have little contribution to the BER. To produce an expression for the upper bound on BER with isolated powers of N ,

TABLE IV

INTERLEAVER GAIN EXPONENT OF THE CONVENTIONAL NON-PRECODED TE VS THE TE-EPCC, $d_E^2 = 7$.

$d_E^2 = 7$	Error pattern classes	TE	TE-EPCC
$m = 4$ $\mu = 1$ $d = 4 \rightarrow d_T$		N^{-1} (4)	$[N^{-1}]_{m_c=3}$ (4) $[N^{-8}]_{m_c=4}$ (6237000)
$d_E^2 = 7$	Error pattern classes	TE	TE-EPCC
$m = 2$ $\mu = 1$ $d = 3 \rightarrow d_T$		N^{-2} (6)	N^{-2} (6)
$d_E^2 = 7$	Error pattern classes	TE	TE-EPCC
$m = 2$ $\mu = 1$ $d = 3 \rightarrow d_T$		N^{-2} (6)	N^{-2} (6)

and at the same time preserve it as an upper bound, we replace the binomial in the denominator by the lower bound [10]:

$$\binom{N}{d} > \frac{(N-d+1)^d}{d!} \simeq \frac{N^d}{d!}.$$

Moreover, to replace the binomial in the numerator with an upper bound that is also a power of $N-d+1$, we first express it as:

$$\binom{N-d}{m-\mu} = \frac{m-\mu+1}{N-d+1} \binom{N-d+1}{m-\mu+1}$$

and employ the upper bound [10]:

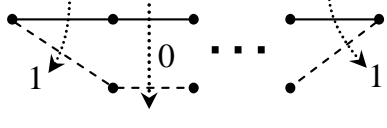

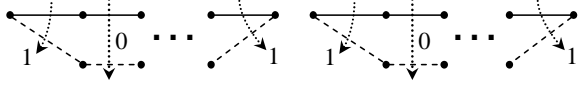

$$\binom{N-d+1}{m-\mu+1} < \frac{(N-d+1)^{m-\mu+1}}{(m-\mu+1)!} \simeq \frac{N^{m-\mu+1}}{(m-\mu+1)!}.$$

These bounds are tight when N is large, and $d, m \ll N$, which holds true in our case. Also we can upper bound the Q function by:

$$Q\left(\frac{d_E}{\sigma}\right) \leq \frac{1}{2} e^{-\frac{d_E^2}{2\sigma^2}}.$$

TABLE V

INTERLEAVER GAIN EXPONENT OF THE CONVENTIONAL PRECODED TE VS THE TE-EPCC, $d_E^2 = \{2, 3, 4, 5\}$.

$d_E^2 = 2$	Error pattern classes	precoded TE	TE-EPCC
$m = 1$ $\mu = 0$ $d^H = 2$ $L = 2$		N^{-1} (2)	N^{-10} $\left(\frac{155925}{4}\right)$
$d_E^2 = 3$	Error pattern classes	precoded TE	TE-EPCC
$m = 2$ $\mu = 1$ $d^H = 3$ $L = 3$		N^{-2} (6)	N^{-10} (779625)
$d_E^2 = 4$	Error pattern classes	precoded TE	TE-EPCC
$m = 2$ $\mu = 0$ $d^H = 4$ $L = 4$		N^{-2} (12)	N^{-9} $\left(\frac{779625}{2}\right)$
$d_E^2 = 5$	Error pattern classes	precoded TE	TE-EPCC
$m = 3$ $\mu = 1$ $d^H = 5$ $L = 5$		N^{-3} (60)	N^{-2} (1)

Substituting these approximate bounds in the BER upper bound in (19), we get a looser but insightful bound:

$$P_b < \frac{1}{2K} \sum_{d_E=1}^{\infty} \sum_{d=2}^{d_T} \sum_{\mu=0}^1 \sum_{m=1}^d \sum_{m: d_E^2 - 2m + \mu = 0 \pmod{4}} \mathbf{B}_{d_E, d, m, \mu} N^{m-\mu-d} e^{-\frac{d_E^2}{2\sigma^2}} \quad (20)$$

where $\mathbf{B}_{d_E, d, m, \mu}$ is given by:

$$\mathbf{B}_{d_E, d, m, \mu} = \mathbf{A}^o(d) \overline{\mathbf{A}}^i(d) \frac{d!}{(m-\mu)!} \left(\frac{1}{2}\right)^{d-m} \binom{d-m}{\frac{d_E^2 - 2m + \mu}{4}} \binom{d-1}{m-1} \quad (21)$$

For the sake of mathematical tractability, we study the interleaver gain exponent of $L_c = 1$ TE-EPCC, i.e. single EPCC subcode per interleave. Utilizing the same approximations as above in the BER bound of TE-EPCC for $L_c = 1$ we get the expression:

$$P_b < \frac{1}{2K} \sum_{d_E=1}^{\infty} e^{-\frac{d_E^2}{2\sigma^2}} \sum_{\mu=0}^1 \left[\sum_{d=2}^{d_T} \sum_{\substack{m=1 \\ m: d_E^2 - 2m + \mu = 0 \bmod 4}}^d \mathbf{B}_{d_E, d, m, \mu} N^{m-\mu-d} \right. \\ \left. - \sum_{d=2}^{\min(d_T, d_c)} \sum_{\substack{m=1 \\ m: d_E^2 = 2m - \mu}}^{\min(d, m_c)} \mathbf{B}_{d_E, d, m, \mu} N^{m-\mu-d} \right]. \quad (22)$$

The expression in (22) is just the expression in (20) with those terms that are correctable by EPCC subtracted. By identifying the maximum exponent of the interleaver length N in (22) and (20), we can compare the asymptotic BER of TE and TE-EPCC in the limit of large interleaver size. Assuming the minimum Hamming weight of the outer RSCC code is 2, we list the maximum interleaver gain exponent per d_E^2 , for TE and TE-EPCC ($d_c = 10$, $m_c = \{3, 4\}$, $L_c = 1$) in Table II for $d_E^2 = \{1, \dots, 4\}$, in Table III for $d_E^2 = \{5, 6\}$, and in Table IV for $d_E^2 = 7$. We also list for each d_E^2 , the generating error patterns and their corresponding parameters d , m , and μ . In addition, under each interleaver gain exponent, we list in parenthesis the corresponding multiplicative coefficient $\mathbf{B}_{d_E, d, m, \mu}$, excluding the term $\mathbf{A}^o(d) \overline{\mathbf{A}}^i(d)$ relating to the outer RSCC Hamming error weight distribution.

For the precoded TE, using the same approximations as above, it can be shown that the BER bound of Appendix A is dominated by the terms:

$$P_b < \frac{1}{2K} \sum_{d_E=2}^{\infty} \sum_{d=2}^N \mathbf{B}_d N^{-\lceil \frac{d}{2} \rceil} e^{-\frac{d_E^2}{2\sigma^2}} \quad (23)$$

where \mathbf{B}_d is given by:

$$\mathbf{B}_d = \mathbf{A}^o(d) \overline{\mathbf{A}}^i(d) \lceil \frac{d}{2} \rceil \binom{d}{\lceil \frac{d}{2} \rceil}. \quad (24)$$

In the derivation of the above bound we only kept terms whose error length L is equal to the Hamming distance d^H since they have the dominant interleaver gain exponent at each d^H . In this case, the dominant error will have $d_E^2 = d^H$ as shown in Table V.

First, we note that for the non-precoded TE, the interleaver gain exponents are all negative for $d_E^2 = 1$ to $d_E^2 = 3$, which are the terms that dominate the BER for medium to high SNRs. Second, we note that the error patterns, for this same range of error Euclidean distances, up to $d = 10$, all belong to the dominant error class. As a result, the TE-EPCC manages to substantially decrease the interleaver gain exponent by a factor of N^9 . Also, for $d_E^2 = 4$, where the TE does not achieve any interleaver gain, the TE-EPCC has an impressive interleaver gain exponent of N^{-9} .

The extremely low exponents suggest that the TE-EPCC will have large gain even for relatively short interleavers, and would thus deliver satisfactory gain for short to medium RSCC codeword sizes. At the same time, for such short interleavers, the TE would considerably suffer in terms of turbo gain. These conclusions will be numerically demonstrated in the next section by evaluating the BER bound for interleavers as short as 100 bits. Furthermore, although $\mathbf{B}_{d_E, d, m, \mu}$ is significantly larger in the TE-EPCC compared to the TE for the same d_E^2 , the term $\mathbf{B}_{d_E, d, m, \mu} N^{m-\mu-d}$ is still several orders of magnitude lower for the TE-EPCC.

Although less important, we also show the interleaver gain for higher error Euclidean distances in Table III and Table IV. Most notably, the TE-EPCC ($d_c = 10$, $m_c = 3$, $L_c = 1$) corrects errors belonging to the dominant error class for $d_E^2 = 5$ and $d_E^2 = 6$, lowering, in the process, the maximum interleaver gain exponent by a factor of N , a turbo gain that becomes more substantial for large interleavers. Actually, for $d_E^2 = 6$, the TE possess no interleave gain, while TE-EPCC BER is dominated by the non-targeted set of non-dominant errors that result in the exponent N^{-1} , still achieving an interleaver gain. On the other hand, the TE-EPCC ($d_c = 10$, $m_c = 3$, $L_c = 1$) would offer no advantage when $d_E^2 = 7$. Note that although all errors belong to χ^{dom} when $m = 4$, their multiplicity m exceeds the maximum multiple-error-pattern correction capability of $m_c = 3$. However, the TE-EPCC ($d_c = 10$, $m_c = 4$, $L_c = 1$) manages to reduce the maximum interleaver gain exponent to N^{-2} , by reducing the contribution of χ^{dom} to $d_E^2 = 7$ by a factor of N^7 .

Comparing the interleaver gain exponents of the TE-EPCC and the precoded TE in Table V, we note that the TE-EPCC focuses on error events of $d_E^2 \leq 4$ and $d^H \leq 4$, reducing the interleaver

exponent by a factor of N^9 , N^8 , and N^7 at $d_E^2 = \{2, 3, 4\}$, respectively, compared to the precoded TE. However, the precoded TE's interleaver exponent is lower by a factor of N compared to the TE-EPCC at $d_E^2 = 5$. Hence, we predict that the TE-EPCC's BER floor will be far lower than that of the precoded TE, while the precoded TE waterfall BER may still be lower owing to the lower interleaver gain exponents of higher Euclidean-distance error events. In summary, the EPCC shapes the error weight spectrum to improve the error floor while preserving gains achieved in the waterfall region. The TE-EPCC is thus a novel turbo equalization approach that enhances the spectrum thinning at low weights, where the error floor is not a strong function of interleaver size as the waterfall region is. Actually, all ML type bounds, including our bound, study the error floor rather than the waterfall region. This works in our favor since our EPCC code's advantage lies there. In addition, our EPCC works in a probabilistic fashion to enhance the minimum Euclidean distance, in addition to the interleaver gain, compared to trellis constrained methods that directly increase the minimum Euclidean distance.

VI. A SISO DECODER FOR EPCC

We have thus far presented the ideal behavior expected from a perfect EPCC decoder. We now discuss a practical implementation of EPCC SISO decoding based on the algebraic single-pattern correcting decoder of Section II and the soft side information made available by the channel observations \mathbf{r} and the outer RSCC SISO decoder.

In the decoder flow chart, a decision is first made on whether the hard input of the decoder $\hat{\mathbf{c}}$ contains a single or a multiple error pattern via the syndrome check. If the initial syndrome check indicates either an error free input, or else, a single error pattern with high reliability [15], then the following formula is used to generate the soft decision reliability for the k -th hard bit in the corrected codeword [23]:

$$\lambda_k = \beta^{iter} \times \lambda_{max} \times \hat{d}_k \quad (25)$$

where $\hat{\mathbf{d}}$ is the bipolar representation of the error-free/corrected bit, λ_{max} is a preset value for the maximum reliability at convergence of turbo performance, and the multiplier $\beta^{iter} < 1$ is useful in incorporating the EPCC SISO decoder in the iterative loop. Note that in an iterative system the level of confidence in bit decisions is lower at the initial iterations, and thus multiplying the generated log likelihood ratios by the back-off factor β^{iter} reduces the risk of error propagation.

On the other hand, if the computed syndrome is unrecognized, signaling a multiple occurrence, then the list decoder is activated which involves computing correlator-based reliability estimates for local dominant patterns in the ML word, as will be explained later in this section. Simulations show that the aforementioned strategy of moving between list-decoding and algebraic single pattern decoding results in improved performance compared to running list-decoding all the time, since at later turbo iterations single error-pattern occurrences are more likely, and syndrome-decoding is more robust in such scenarios.

In communicating with the other building blocks of the turbo system, the EPCC decoder receives the interleaved extrinsic LLR λ coming from the outer RSCC code as an *a priori* input in calculating its error pattern *a posteriori* probabilities. On the other hand, in the final soft output stage, after generating a list of the most probable candidate codewords and their likelihoods, the decoder uses the list to calculate the output bit-level decision reliabilities that serve as the *a priori* input LLR to the outer RSCC SISO decoder.

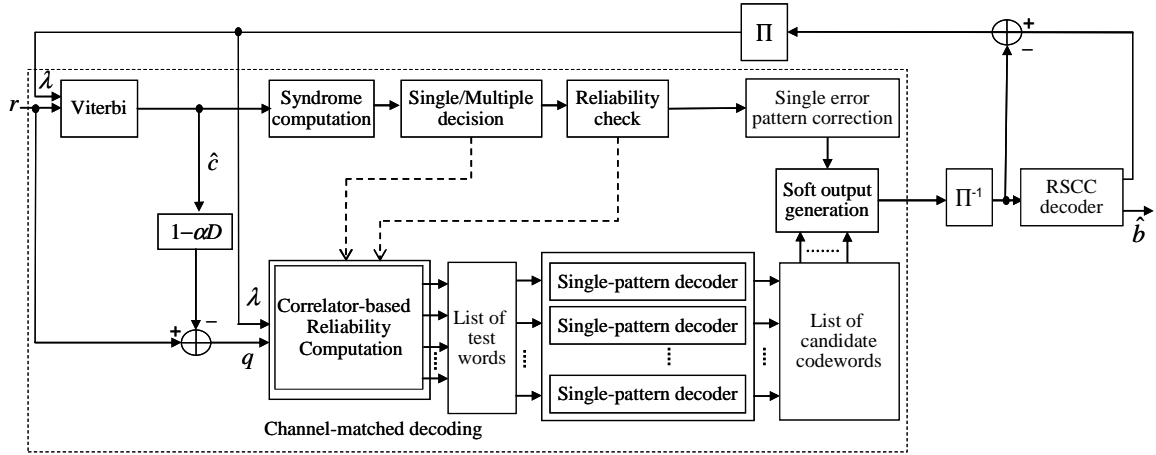


Fig. 5. The TE-EPCC block diagram.

The internal workings of the EPCC soft output list decoder considered here consists of four stages, see Fig. 5:

- The probability of a dominant single error event is estimated at each likely starting position.
- The test error word list is generated by inserting the most probable combination of dominant error patterns into the channel detector ML output.
- An array of parallel single-pattern correcting decoders decode the test words to produce a

list of valid codewords and accompanying likelihoods.

- The list of candidate codeword likelihoods is used to generate bit-level decisions along with their reliabilities.

A. Dominant Error Probability Estimation

At this decoder stage we estimate the likelihood $C(e_k^{(i)})$ at all possible locations k that the dominant error event $e^{(i)}$ can occur. A bank of error event correlators can be used for this purpose as was shown in [15] and [16]. Let r_k be the channel detector input sequence $r_k = c_k * h_k + w_k$, where c_k is the bipolar representation of the transmitted codeword sequence, h_k is the channel response of length l_h , and w_k is zero-mean AWGN noise with variance σ^2 . Also, let $q_k = r_k - (\hat{c}_k * h_k) = (c_k - \hat{c}_k) * h_k + w_k$ be the channel detector's output error sequence. If a target error pattern sequence $e_k^{(i)}$ occurs at positions from $k = j$ to $k = j + l_i - 1$, then q_k can be written as

$$\begin{aligned} q_k &= [\mathbf{c} - \hat{\mathbf{c}}]_j^{j+l_i-1} * h_k + w_k \\ &= [\mathbf{e}^{(i)}]_j^{j+l_i-1} * h_k + w_k \\ &= [\xi^{(i)}]_j^{j+l_i^h} + w_k \end{aligned} \quad (26)$$

where $\xi_k^{(i)}$ is the channel response of the error sequence, and is given by $\xi_k^{(i)} = e_k^{(i)} * h_k$, and $l_i^h = l_i + l_h - 2$. Using the MAP criterion, we can derive an estimate of the likelihood of $e_k^{(i)}$ by measuring the distance of the resulting $\mathbf{q}(e_k^{(i)})$ to the channel observation \mathbf{r} relative to the error free sequence $\mathbf{q}(\hat{\mathbf{c}})$, which simplifies to just the difference in Euclidean distances of $\mathbf{q}(e_k^{(i)})$ and $\mathbf{q}(\hat{\mathbf{c}})$ measured to \mathbf{r} in the ML sense. The ML postulate becomes useful if $\hat{\mathbf{c}}$ is assumed i.i.d, or when it is the best that can be done when no *a priori* side information is available. In practice, though, the side information in the form of bit-level log-likelihood ratios λ_k can be efficiently provided by the outer constituent code in the turbo sense. For each $e_k^{(i)}$, we then estimate [15]:

$$\begin{aligned} C(e_j^{(i)}) &= \sum_{k=j}^{j+l_i^h} \frac{1}{2\sigma^2} \left(q_k^2 - (q_k - \xi_k^{(i)})^2 \right) \\ &\quad - \left(\sum_{k=j, \hat{c}_k=+1}^{j+l_i^h} \lambda_k - \sum_{k=j, \hat{c}_k=-1}^{j+l_i^h} \lambda_k \right) \end{aligned} \quad (27)$$

where λ_k is the *a priori* LLR of the error-event bit at position k as received from the outer soft decoder, and we are assuming here that error event sequences do not include 0 bits, i.e., the ML

sequence and error sequence do not agree for the entire duration of the error event. Finally, the expanded list of dominant errors and their likely positions is sorted according to the computed reliabilities.

B. Generation of the Test Error Word List

In order to expand the decoding-sphere radius of the algebraic single-error correcting code, and to benefit from channel side information, we adopt a list-decoding structure that resembles Chase decoding [24] in the sense of generating test vectors at the parallel decoder input. However, a pivotal difference in the methodologies is in the test word construction stage of the EPCC decoder, where we flip multi-bit dominant error patterns, rather than individual independent bits as in Chase decoding. The resulting list of test vectors may contain one or more words that are just one dominant error pattern away from the correct codeword in terms of Hamming distance. Hence, if the resulting set of test error words is decoded by an array of single-pattern correcting decoders, then, one or more codewords in the list of valid candidate codewords can be the correct codeword with high probability. This novel pattern-level extension of the Chase decoding algorithm was first proposed in [16] in the context of SISO decoding of EPCC as a building block in TE, and in [15] in the context of list decoding of algebraic single-pattern correcting EPCC. Recently, a pattern-flipping Chase was also studied in [25]. This later approach differs from the earlier work in [15], [16] only in the block that estimates pattern reliabilities, where SOVA is utilized to estimate error event probabilities instead of a bank of error-event-matched correlators.

The probability measure of a given test word with a particular combination of dominant error patterns is the product of the probabilities of the constituent errors. In the construction of test words, we select the most probable errors in the sense of maximizing the correlator function of (27).

The requirement to have m_c -error-pattern-correction capability using the single-pattern correcting decoders, dictates that test words must include up to $m_c - 1$ single dominant error patterns. Starting from the l_{max} most probable such dominant errors, one can think of $\sum_{j=1}^{m_c-1} \binom{l_{max}}{j}$ ways of corrupting the ML word with up to $m_c - 1$ local error patterns. From this large set of potential combinations, a relatively small subset of most probable combinations needs to be chosen to maintain reasonable complexity. One can think of many different ways of effectively constructing

such a list [15] based on the probable local error patterns that have been identified.

C. Parallel Algebraic Decoding

The list of test error words generated above is delivered to an array of single-error-pattern correcting decoders that work in parallel to generate the candidate codeword list. The number of parallel decoders is identical to the size of the test word list, and is a crucial parameter that controls the EPCC decoder's complexity/performance tradeoff.

D. Generation of Soft Output

The candidate codeword list constructed by our “pattern-level” list decoder is used to calculate the more familiar bit-level reliabilities that constitute the output soft information supplied by the EPCC SISO decoder. We measure the probability of a candidate codeword given the observed word by the product of the probabilities of each “local” error pattern forming the candidate word. Specifically, let \mathbf{c} represent a candidate codeword with, say, K error-pattern corruption with respect to the ML word $\hat{\mathbf{c}}$. Then, the *a posteriori* probability of this particular test word, $Pr(\mathbf{c}|\hat{\mathbf{c}}, \mathbf{r})$, is estimated by multiplying the probability estimates of the K local patterns, given the channel observation \mathbf{r} at the detector input.

Given the list of codewords and their accompanying *a posteriori* probabilities, the reliability λ_k of the coded bit c_k is evaluated as:

$$\lambda_k = \log \frac{\sum_{\mathbf{c} \in \mathbf{S}_k^+} Pr(\mathbf{c}|\hat{\mathbf{c}}, \mathbf{r})}{\sum_{\mathbf{c} \in \mathbf{S}_k^-} Pr(\mathbf{c}|\hat{\mathbf{c}}, \mathbf{r})} \quad (28)$$

where \mathbf{S}_k^+ is the set of candidate codewords where $c_k = +1$, and \mathbf{S}_k^- is the set of candidate codewords where $c_k = -1$. The quantity in (28) is utilized when the candidate codewords do not all agree on the bit decision for location k . In the event that all codewords do agree on the decision for c_k , a method used by [23] is adopted for generating soft information as follows

$$\lambda_k = \beta^{iter} \times \lambda_{max} \times \hat{d}_k \quad (29)$$

where \hat{d}_k is the bipolar representation of the agreed-upon decision, and λ_{max} and $\beta^{iter} < 1$ have the same function as in (25).

VII. NUMERICAL ANALYSIS AND SIMULATION RESULTS

Utilizing the analytic approximation of the BER of conventional TE and TE-EPCC systems, we study the relative performance of these systems for different levels of the severity of ISI. We also study the special case of the dicode channel $1 - D$ in a variety of channel conditions. We will assume throughout the analysis that the SNR rate penalty (in dB) is proportional to $10 \log_{10} \frac{1}{R}$, where R is the code rate.

A. BER-Bound Validation for the Dicode Channel

The log of the average Euclidean distance distribution of the dicode channel, $\log \bar{T}(d_E)$, is shown in Fig. 6 for conventional TE and TE-EPCC systems. Fig. 6 also includes the Euclidean error distribution for the precoded TE, derived in a similar way to that of [11]. $\log \bar{T}(d_E)$ is calculated for a TE with $K = 4096$, a rate $1/2$ base constituent RSCC, punctured to rate $R = 8/9$ with generator polynomial connections (31, 33) in octal format, and $L_c = 7$ EPCC with $m_c = 3$ and $d_c = 10$. Each EPCC subcode is a (630, 616) systematic cyclic code of rate 0.98 shortened to accommodate $L_c = 7$ subcodes per interleave. From the average Euclidean distance distribution,

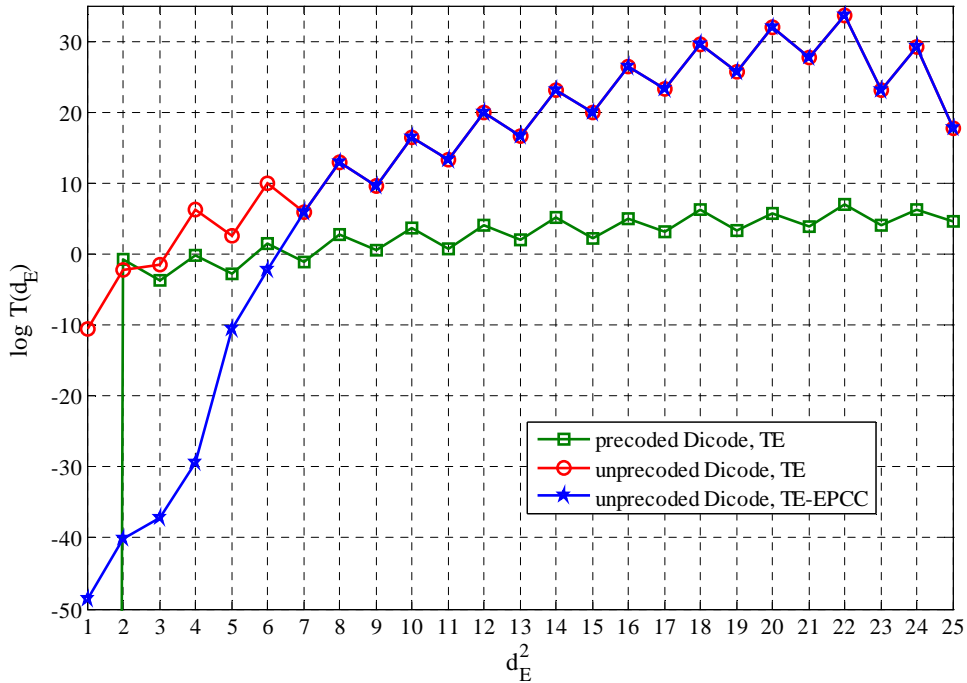


Fig. 6. $\log \bar{T}(d_E)$ for various TE systems, and (31, 33) RSCC.

we can conclude that the precoded TE exhibits larger interleaver gains compared to the non-precoded TE in the waterfall region, i.e. low to medium SNRs. This is because $\log \bar{T}(d_E)$ is lower for the precoded TE everywhere when $d_E^2 > 3$. However, for higher SNRs, at the error floor region, the contribution of squared Euclidean distance 2 becomes stronger, and as seen in the figure, the average number of Hamming weight 2 errors that generate $d_E^2 = 2$ is more for the precoded compared to the non-precoded decode channel. On the other hand, the EPCC concentrates on low Euclidean distances, reducing their frequency substantially up to $d_E^2 = 6$. This results in improved BER in the error floor and yields a similar waterfall BER compared to the conventional unprecoded TE.

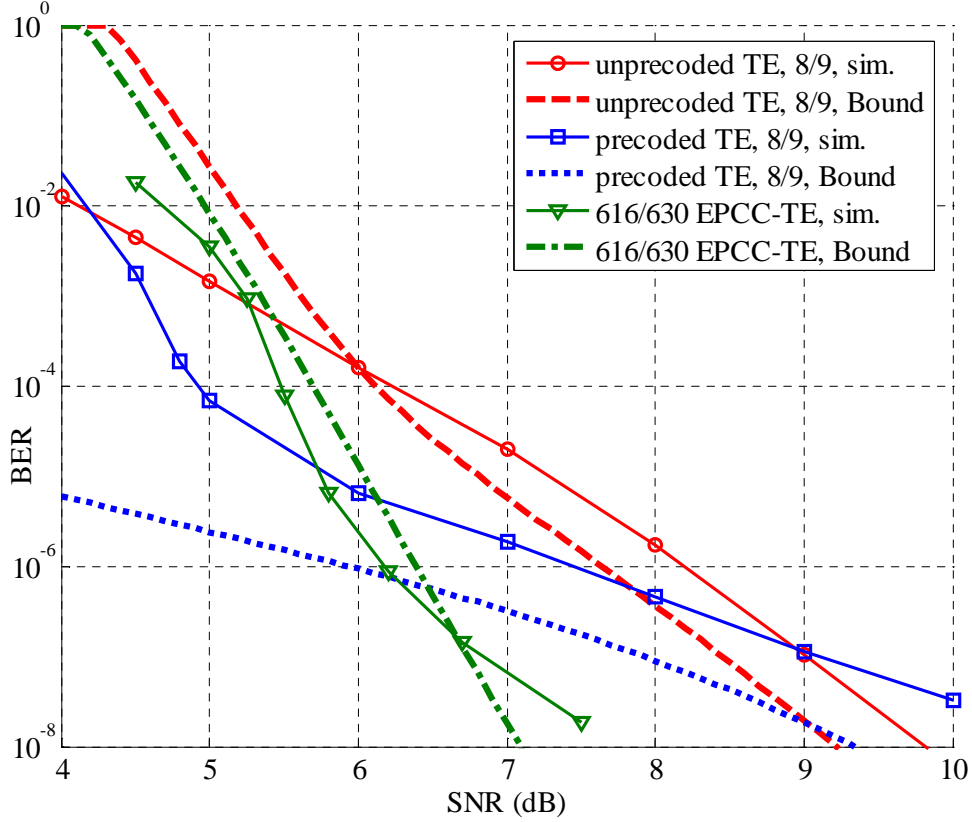


Fig. 7. Simulated vs bounded BER for various TE systems using (31, 33) RSCC.

In Fig. 7, the simulated BER is shown for conventional TEs with precoded and non-precoded decode channels and the TE-EPCC with a non-precoded decode channel. Moreover, the simulated BER is compared with the estimated BER bound computed for the same parameters as in Fig. 6. The TE-EPCC is decoded via the practical soft decoder described earlier in Section VI, where

we implement 5 turbo iterations of the non-precoded TE and 10 turbo iterations of precoded TE and TE-EPCC systems; we used up to 100 test patterns in the list decoder of the TE-EPCC. The number of turbo iterations is chosen for each system such that the turbo gain saturates. The figure shows that the TE-EPCC has definite performance advantage in the low error rate region. The actual simulation curve comes above the analytical bound for the TE-EPCC at very low BERs. This arises from imperfect uniform interleaving in the practical decoders as also pointed out in [11]. Nonetheless, the actual gain gaps between the TE-EPCC and the conventional TEs seem even large than predicted by the bound; this is mainly attributed to the higher sensitivity of conventional TEs to the interleaver design compared to the TE-EPCC, an argument based on the interleaver gain exponent of both systems.

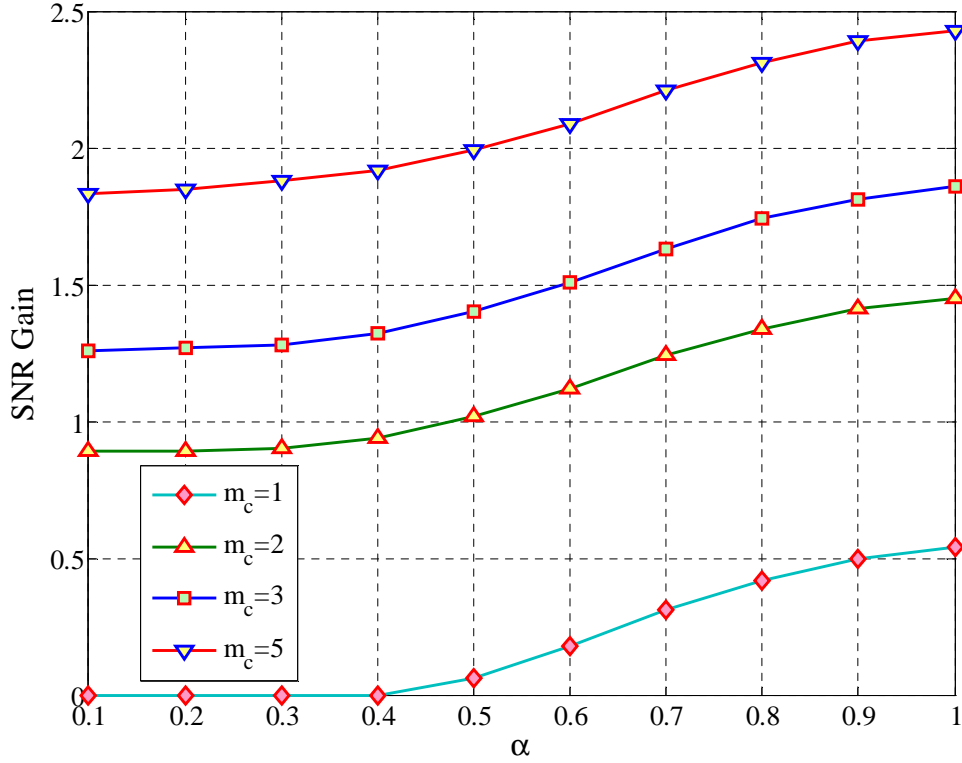


Fig. 8. SNR gain (dB) of the TE-EPCC over the non-precoded TE at a BER of 10^{-7} as function of ISI severity level α for outer RSCC (7, 5), punctured $R = 8/9$, and interleaver size $N = 616$.

B. The Severity of ISI

In Fig. 8 we plot the SNR gain of the TE-EPCC over the non-precoded TE defined as the difference in the minimum SNR required to achieve a BER of 10^{-7} for both systems. The BER of the non-precoded TE improves as $\alpha \rightarrow 0$, since the Euclidean distance of dominant Hamming errors grows with their Hamming distance when $\alpha < 1$, where the error length is linearly proportional to $(1 - \alpha^2) > 0$. On the other hand, for a given EPCC correction power m_c , the BER of the TE-EPCC remains almost the same as $\alpha \rightarrow 0$ since dominant Hamming errors are correctable irrespective of their Euclidean weight. The net result is a higher SNR gain furnished by the TE-EPCC as $\alpha \rightarrow 1$. Furthermore, for a given α , the SNR gain of the TE-EPCC grows as m_c is increased, where a 2 dB improvement can be achieved by increasing m_c from 1 to 5 for all the range of α .

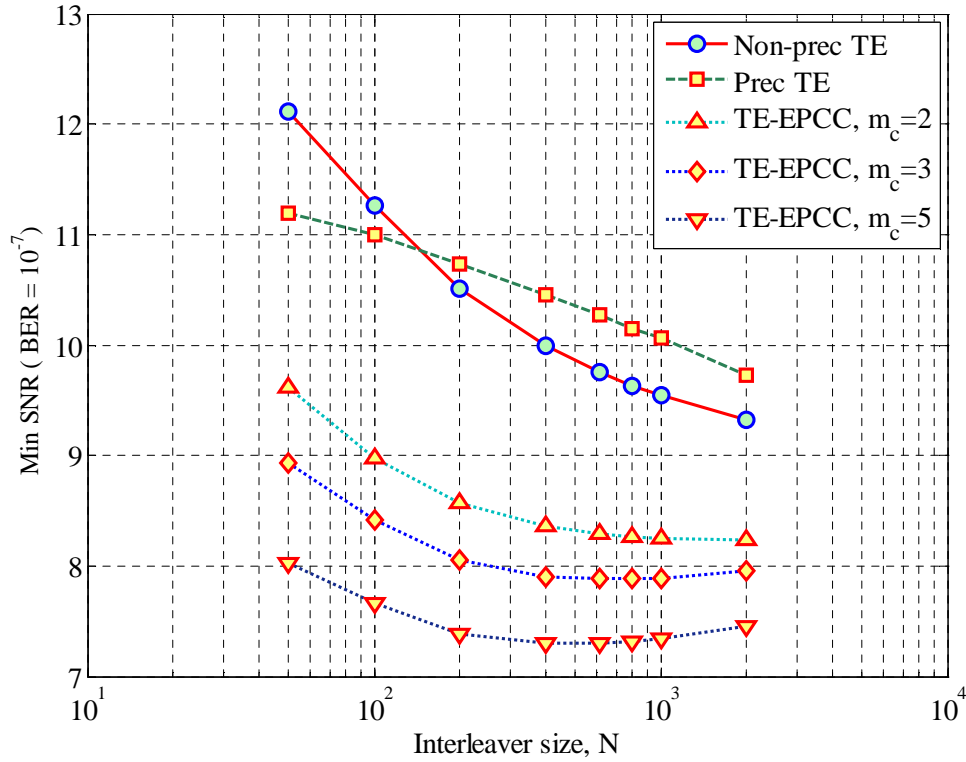


Fig. 9. Minimum SNR (dB) required to achieve a BER of 10^{-7} using (7, 5) RSCC, punctured rate 8/9, TE-EPCC $\{m_c = 3, d_c = 10, L_c = 1\}$, and different interleaver sizes.

C. Interleaver Gain

In Fig. 9 we compare the minimum SNR to achieve a BER of 10^{-7} . The curves of the conventional TE with the precoded and non-precoded dicode channels and the TE-EPCC for the non-precoded dicode channel are shown for interleaver lengths $N = \{50, 100, 200, 400, 600, 800, 1000, 2000\}$ bits, punctured rate $R = 8/9$ RSCC with connections $(7, 5)$ in octal format, and EPCC with $m_c = \{2, 3, 5\}$ and $d_c = 10$. The SNR gain of the TE-EPCC over the non-precoded TE using $N = 50$ is 2.3 dB, 3 dB, and 4 dB for EPCC correction powers $m_c = \{2, 3, 5\}$, respectively. On the other hand, using $N = 2000$, this shrinks to 1.1 dB, 1.3 dB, and 1.8 dB for EPCC correction powers $m_c = \{2, 3, 5\}$, respectively.

We note that as the interleaver size N of the TE-EPCC increases, the turbo gain of TE-EPCC increases accordingly. Also, since we maintain the same number of parity bits as the codeword length increases, less SNR rate penalty is incurred as N increases. On the other hand, the probability of $m > 1$ multiple errors per subcode increases for larger N , surpassing EPCC's correction capability. Due to these conflicting effects of the TE-EPCC, its minimum SNR plateaus and even increases as N increase beyond a certain point. All in all, the relative advantage of the TE-EPCC in practical system seems most visible with small interleaver sizes. Furthermore, as can be observed in the figure, increasing m_c is also most effective for smaller interleaver sizes.

In practical EPCC code construction, in order to obtain shorter EPCC code lengths, while serially concatenating one EPCC subcode per RSCC interleave, i.e. $L_c = 1$, the EPCC code length is shortened from the long $(630, 616)$ EPCC at the same level of redundancy. While to support interleaver sizes above 630, we duplicate EPCC subcodes, i.e. $L_c > 1$, and use shortening to fit fractions of EPCC subcodes in one interleave. For instance, we implement $(114, 100)$ EPCC of rate 0.88 for interleaver length $N = 100$, and long $(630, 616)$ EPCC + shortened $(398, 384)$ EPCC for $N = 1000$.

D. SNR Gain as Function of L_c and m_c

The performance of TE-EPCC can be further improved by increasing its multiple error correction capability m_c , per subcode. However, the complexity of the practical decoder would increase accordingly as more test words have to be constructed in the list decoder. Fig. 10 shows TE-EPCC's SNR gain over the non-precoded TE for several BER operating points, $N = 1200$, punctured $R = 8/9$ RSCC with connections $(7, 5)$, and $L_c = 1$ EPCC with maximum correction

capability increased from $m_c = 1$ to $m_c = 10$ and $d_c = 10$. The curves demonstrate that TE-EPCC's SNR gain grows almost linearly as m_c is increased. Another design method to increase

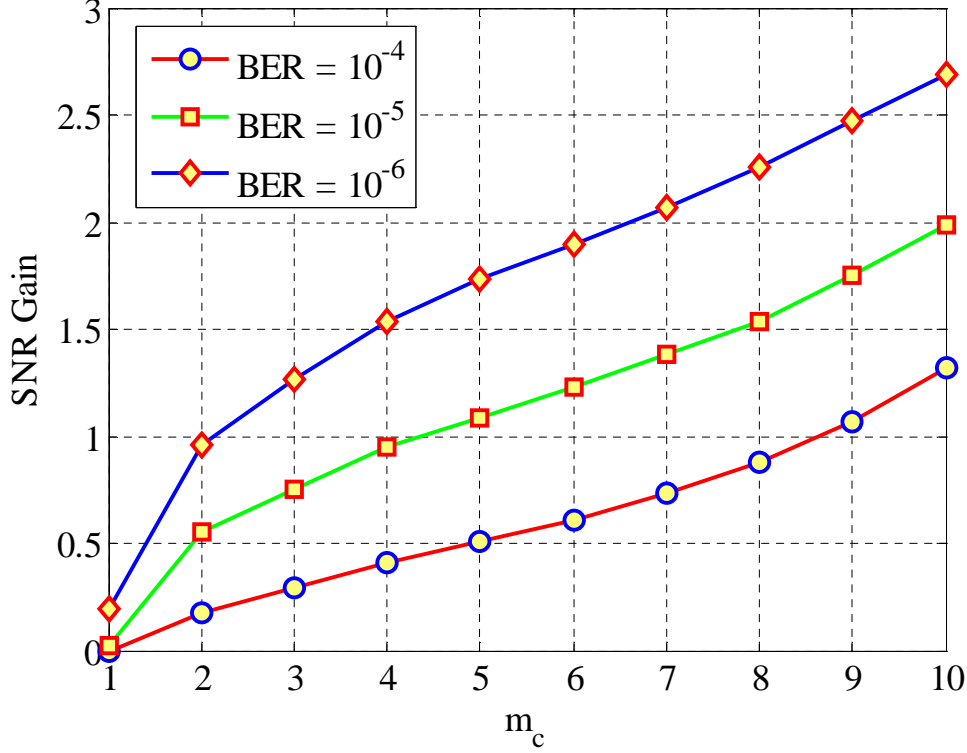


Fig. 10. SNR gain of the TE-EPCC over the non-precoded TE at several BER operating points as function of m_c for the dicode channel, outer RSCC (7, 5), punctured $R = 8/9$, and interleaver size $N = 616$.

the correction capability of TE-EPCC, without considerably increasing its complexity, is to use $L_c > 1$ EPCC subcodes per interleave. To study the design space spanned by m_c and L_c for a given interleave, we evaluate the BER bound for the set composed of the Cartesian product of the sets $m_c = \{1, 2, 3, 4\}$ and $L_c = \{2, 3, 4, 5, 6\}$. Then we plot a continuous contour of the minimum SNR to achieve $\text{BER}=10^{-5}$ by interpolating the values found at the elements of the Cartesian product. A contour plot with an SNR step of 0.1 dB is shown in Fig. 11 for $N = 1200$, punctured-rate 8/9 RSCC with connections (7, 5), and EPCC with different combinations of m_c and L_c , and $d_c = 10$. We note that the combinations $\{L_c = 2, m_c = 4\}$ and $\{L_c = 5, m_c = 3\}$ require a similar minimum SNR=6.6 dB to achieve a floor BER of 10^{-5} . Nonetheless, as can be seen in Fig. 11 the slope the equi-SNR contour lines decreases for higher L_c and lower m_c . This means that as the number of subcodes L_c increases per interleave, the correction capability

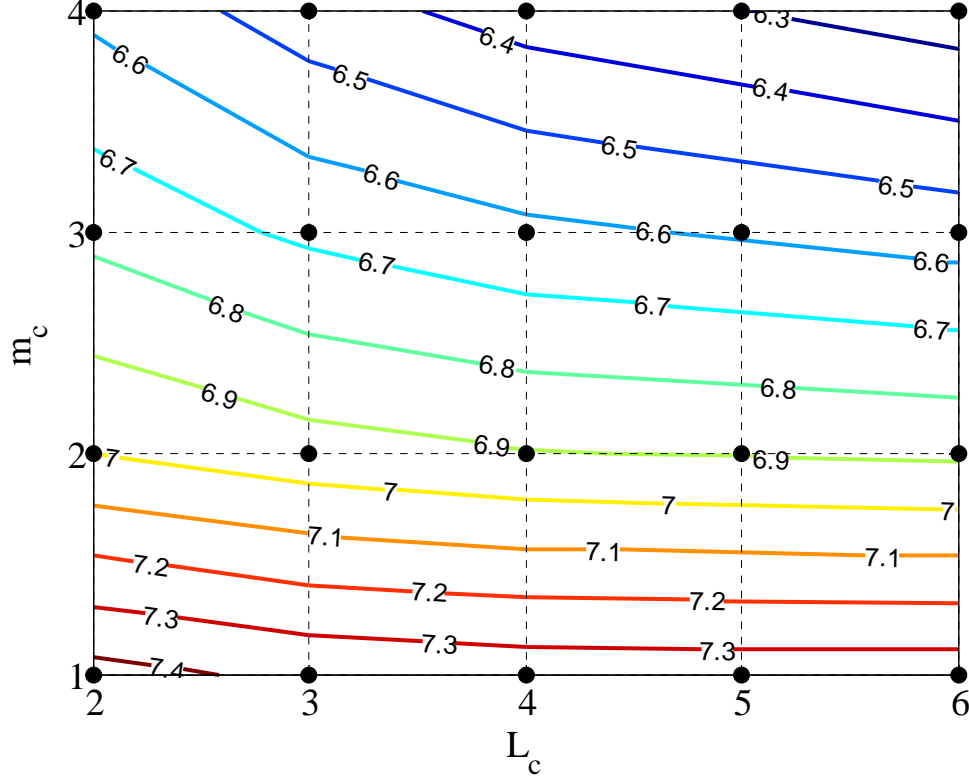


Fig. 11. An interpolated contour plot of the minimum SNR required by TE-EPCC to achieve $\text{BER}=10^{-5}$ for different combinations of m_c and L_c , $N=1200$, and RSCC (7, 5) of punctured $R = 8/9$.

plateaus, especially when $m_c \leq 2$. This is due to the higher level of redundancy required for shortened EPCC to maintain the maximum correction capability m_c of longer EPCC. For instance, at one extreme, to maintain the correction capability at a shortened EPCC code length of 44 bits, i.e. $L_c = 40$ and $N = 1200$, a shortened EPCC of rate 0.68 would incur a staggering rate penalty of 1.7 dB. An alternative concatenation approach that avoids the rate penalty of serial concatenation to a short inner EPCC is discussed in [26].

E. Puncturing Rate

We also wish to study TE-EPCC advantage at various total system rates and distributions of redundancy between the outer RSCC and inner EPCC subcode. In Fig. 12, we compare the simulated BER of the conventional non-precoded TE and the TE-EPCC, for interleaver length $N = 4312$, different punctured-rate RSCC with connections (7, 5), and EPCC with $m_c = 3$ and $d_c = 10$. The results show that the TE-EPCCs composed of either $L_c = 7$ (630, 616) EPCC or

$L_c = 22$ (210, 199) EPCC concatenated to rate $\frac{5}{6}$ TE, achieve the same BER in the error floor region. Furthermore, they both outperform comparable rate conventional TEs, with $L_c = 22$ TE-EPCC furnishing a gain of 1.5 dB with respect to rate $\frac{3}{4}$ TE at $\text{BER}=10^{-6}$, and $L_c = 7$ TE-EPCC delivering similar gain over rate $\frac{5}{6}$ TE. Moreover, either TE-EPCCs deliver 1 dB SNR gain over the precoded TE of punctured-rate $\frac{5}{6}$. For a complete investigation of a wide range of coding rates, we plot the minimum SNR required to achieve a BER of 10^{-7} for punctured coding rates from $\frac{2}{3}$ to $\frac{9}{10}$, comparing the conventional non-precoded and precoded TE to the TE-EPCC. Such a comparison is shown in Fig. 13 for interleaver length $N = 1200$, different punctured-rate RSCC with connections (7, 5), and $L_c = 1$ EPCC with $m_c = 3$ and $d_c = 10$. We conclude from the results that the TE-EPCC delivers a uniform gain of 1.5 dB for puncturing rates above $\frac{3}{4}$. The abnormal peak in BER for puncturing rate $\frac{6}{7}$ is due to the particular choice of puncturing table. The reason why the precoded TE outperforms the TE-EPCC for puncturing rates $\frac{2}{3}$ and $\frac{3}{4}$ can be explained by examining $A^o(2)$ for those puncturing rates, where it was shown in [22], using a similar approach to [27], that the outer RSCC does not generate Hamming weight 2 errors for these low puncturing rates. Hence, since the BER performance of precoded TE is dominated by such errors in the floor region, its BER is significantly improved surpassing the TE-EPCC at those rates. In summary, the precoded TE is more effective when the minimum Hamming distance of the outer code is larger than 2. Hence, its less effective for high rate simple punctured codes, where its hard to design punctured rates of this property. Therefore, the TE-EPCC is more effective at high code rates for which simple puncturing is utilized.

VIII. CONCLUSION

In this work, we have studied the BER of the serial concatenation of EPCC and interleaved RSCC over ISI channels as an alternative to a single RSCC with and without an inner precoder. To facilitate the study of system performance for a wide range of coding rates, interleaver sizes, and EPCC design parameters, we have derived an approximate upper bound on the BER of the TE-EPCC that is easy to evaluate and that scales well with system parameters. We have also shown how EPCC enhances TE performance by reducing the frequency of error words of low Euclidean distance, which dominate the BER both in the waterfall and error floor regions. Numerical results, calculated via the derived bound, indicate that the TE-EPCC delivers substantial gain for short interleaver lengths compared to the precoded and non-precoded TE, which makes it more

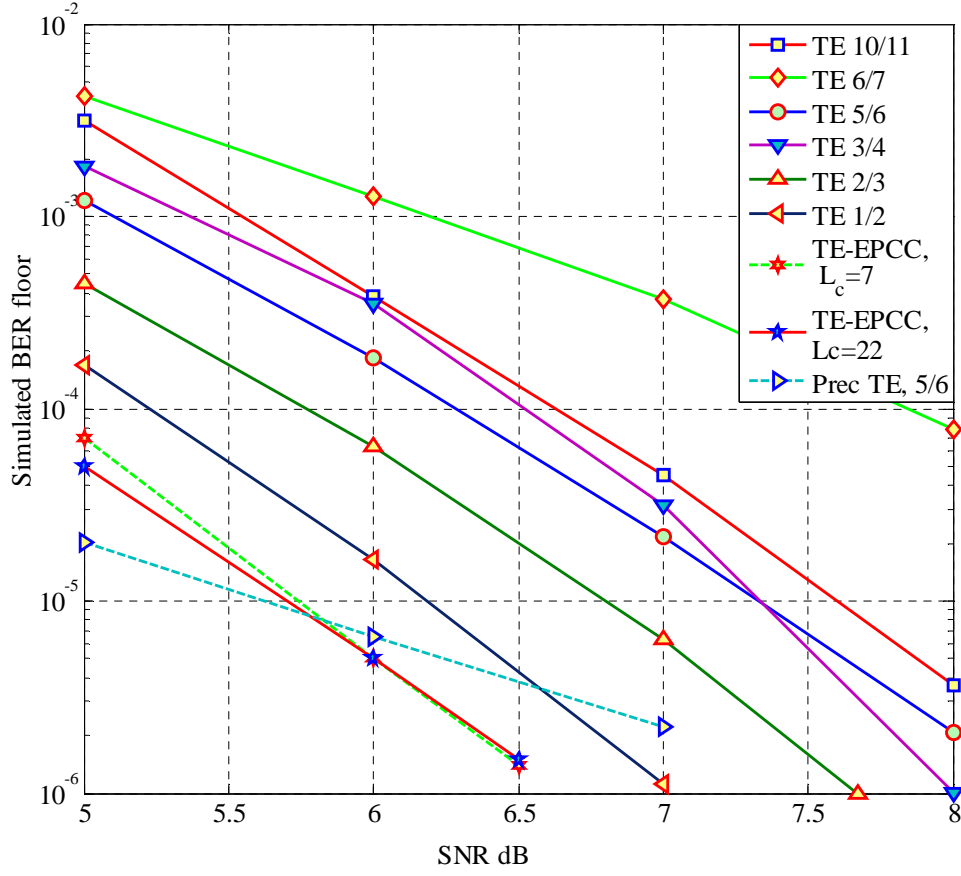


Fig. 12. Simulation of the TE-EPCC and the conventional TE for various rates.

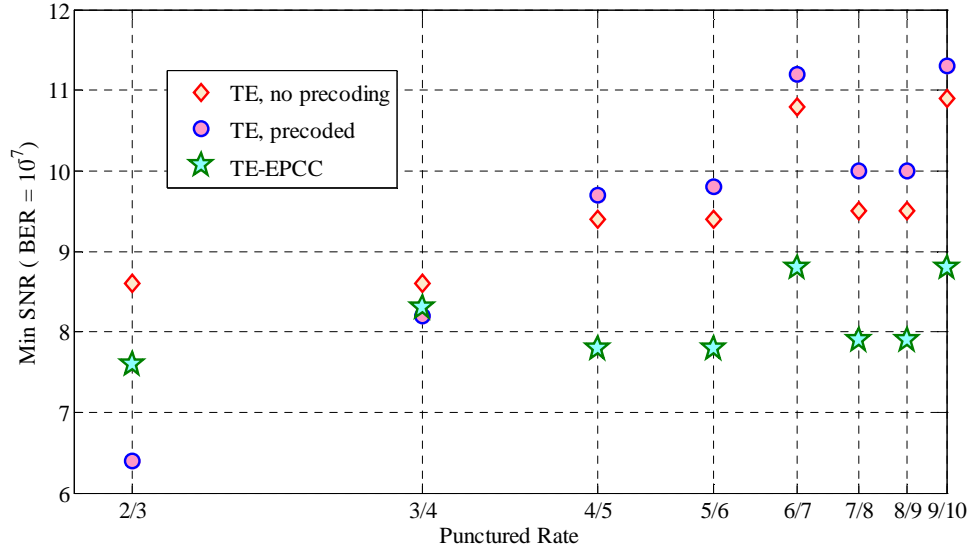


Fig. 13. Minimum SNR required to achieve 10^{-7} BER for various TE systems as function of the outer (7, 5) RSCC punctured rate, $N = 1200$, and $\{m_c = 3, L_c = 1, d_c = 10\}$ TE-EPCC.

attractive than the conventional TE for hardware implementation. Also, we have demonstrated that the TE-EPCC furnishes a uniform gain of 1.5 dB for puncturing rates above 0.75, which makes it suitable for high rate applications, such as magnetic and optical recording applications, while the precoded TE is a better choice for lower coding rates.

APPENDIX

A. Simple BER bound expressions for $L_c = 1$

1) *Non-precoded TE:* As discussed above for the $1 - \alpha D$ channel, the Euclidean distance of a compound error event of multiplicity m and γ crossing branches is given by:

$$d_E^2 = 4\alpha\gamma + (1 - \alpha)^2 d + 2\alpha m - \mu\alpha^2.$$

where d is the Hamming distance of the compound error. In order to compare the channel SNR of different levels of ISI α , we make use of the noise variance normalization:

$$\tilde{\sigma}^2 = \frac{1 + \alpha^2}{2} \sigma^2$$

in which case the decode channel has the base noise variance σ^2 . When the EPCC is turned off, the expression of the TE's BER reduces to:

$$\begin{aligned} P_b &\leq \frac{1}{K} \sum_{d_E=1}^{\infty} Q\left(\frac{d_E}{\tilde{\sigma}}\right) \sum_{d=2}^{d_T} \frac{\mathbf{A}^o(d) \overline{\mathbf{A}}^i(d)}{\binom{N}{d}} \\ &\quad \sum_{\mu=0}^1 \sum_{\substack{m=1 \\ m: \gamma \geq 0, \gamma \in \mathbf{N}}}^d \binom{d-m}{\gamma} \\ &\quad \left(\frac{1}{2}\right)^{d-m} \binom{N-d}{m-\mu} \binom{d-1}{m-1} \\ \gamma &= \frac{d_E^2 - 2\alpha m + \mu\alpha^2 - (1-\alpha)^2 d}{4\alpha} \end{aligned} \tag{30}$$

A good approximation of the Q function that is accurate for a wide range of abscissa is borrowed from [28], and is given by:

$$Q\left(\frac{d_E}{\tilde{\sigma}}\right) \simeq \frac{1}{12} e^{-\frac{d_E^2}{2\tilde{\sigma}^2}} + \frac{1}{4} e^{-\frac{2d_E^2}{3\tilde{\sigma}^2}}.$$

Hence, P_b in (30) is composed of two terms as in:

$$P_b(\tilde{\sigma}) = \frac{1}{12}\check{P}_b(\sqrt{2}\tilde{\sigma}) + \frac{1}{4}\check{P}_b\left(\sqrt{\frac{3}{2}}\tilde{\sigma}\right)$$

where

$$\begin{aligned} \check{P}_b(\tilde{\sigma}) \leq & \frac{1}{K} \sum_{d=2}^{\infty} \mathbf{A}^o(d) \overline{\mathbf{A}}^i(d) \\ & \sum_{\mu=0}^1 \sum_{m=1}^d \sum_{\gamma=0}^{d-m} \binom{d-1}{m-1} \binom{d-m}{\gamma} \left(\frac{1}{2}\right)^{d-m} \\ & N^{m-\mu-d} \frac{d!}{(m-\mu)!} e^{-\frac{4\alpha\gamma+(1-\alpha)^2d+2\alpha m-\mu\alpha^2}{\tilde{\sigma}^2}}. \end{aligned} \quad (31)$$

Evaluating the summation over γ by utilizing the binomial identity we obtain:

$$\begin{aligned} \check{P}_b(\tilde{\sigma}) \leq & \frac{1}{K} \sum_{d=2}^{\infty} \frac{\mathbf{A}^o(d) \overline{\mathbf{A}}^i(d)}{(2N)^d} \\ & \sum_{\mu=0}^1 N^{-\mu} e^{-\frac{(1-\alpha)^2d-\mu\alpha^2}{\tilde{\sigma}^2}} \left(1 + e^{-\frac{4\alpha}{\tilde{\sigma}^2}}\right)^d \\ & \sum_{m=1}^d \frac{d!}{(m-\mu)!} \binom{d-1}{m-1} \left(\frac{2Ne^{-\frac{2\alpha}{\tilde{\sigma}^2}}}{1 + e^{-\frac{4\alpha}{\tilde{\sigma}^2}}}\right)^m. \end{aligned} \quad (32)$$

After some algebraic manipulation, (32) simplifies to:

$$\begin{aligned} \check{P}_b(\tilde{\sigma}) \leq & \frac{1}{K} \sum_{d=2}^{\infty} \frac{\mathbf{A}^o(d) \overline{\mathbf{A}}^i(d)}{(N)^d} \\ & \sum_{\mu=0}^1 N^{1-\mu} e^{-\frac{1+\alpha^2(d-\mu)}{\tilde{\sigma}^2}} \frac{d!}{(1-\mu)!} \left(\cosh\left(\frac{2\alpha}{\tilde{\sigma}^2}\right)\right)^{d-1} \\ & {}_1F_1\left(1-d; 2-\mu; -N \operatorname{sech}\left(\frac{2\alpha}{\tilde{\sigma}^2}\right)\right) \end{aligned} \quad (33)$$

where ${}_1F_1$ is the confluent hypergeometric function of the first kind [29], [30].

2) *TE-EPCC*: The BER of TE-EPCC, P_b^{epcc} , is expressed as the residual error rate after subtracting the error rate component that is correctable by EPCC, P_b^C , from the conventional non-precoded TE BER, P_b , as expressed in:

$$P_b^{epcc}(\tilde{\sigma}) = P_b(\tilde{\sigma}) - P_b^C(\tilde{\sigma}).$$

Similar to (31), using the Q function approximation we have

$$P_b^C(\tilde{\sigma}) = \frac{1}{12}\check{P}_b^C(\sqrt{2}\tilde{\sigma}) + \frac{1}{4}\check{P}_b^C\left(\sqrt{\frac{3}{2}}\tilde{\sigma}\right)$$

with

$$\begin{aligned} \check{P}_b^C(\tilde{\sigma}) \leq & \frac{1}{K} \sum_{d=2}^{\infty} \frac{\mathbf{A}^o(d)\bar{\mathbf{A}}^i(d)}{(2N)^d} \sum_{\mu=0}^1 N^{-\mu} e^{-\frac{(1-\alpha)^2 d - \mu \alpha^2}{\tilde{\sigma}^2}} \\ & \sum_{m=1}^{\min(d, m_c)} \frac{d!}{(m-\mu)!} \binom{d-1}{m-1} \left(2N e^{-\frac{2\alpha}{\tilde{\sigma}^2}}\right)^m \end{aligned} \quad (34)$$

which can be expanded into two sum terms depending on the value of d :

$$\begin{aligned} \check{P}_b^C(\tilde{\sigma}) \leq & \frac{1}{K} \sum_{d=2}^{m_c} \frac{\mathbf{A}^o(d)\bar{\mathbf{A}}^i(d)}{(2N)^d} \sum_{\mu=0}^1 N^{-\mu} e^{-\frac{(1-\alpha)^2 d - \mu \alpha^2}{\tilde{\sigma}^2}} \\ & \sum_{m=1}^d \frac{d!}{(m-\mu)!} \binom{d-1}{m-1} \left(2N e^{-\frac{2\alpha}{\tilde{\sigma}^2}}\right)^m \\ & + \frac{1}{K} \sum_{d=m_c+1}^{\infty} \frac{\mathbf{A}^o(d)\bar{\mathbf{A}}^i(d)}{(2N)^d} \sum_{\mu=0}^1 N^{-\mu} e^{-\frac{(1-\alpha)^2 d - \mu \alpha^2}{\tilde{\sigma}^2}} \\ & \sum_{m=1}^{m_c} \frac{d!}{(m-\mu)!} \binom{d-1}{m-1} \left(2N e^{-\frac{2\alpha}{\tilde{\sigma}^2}}\right)^m. \end{aligned} \quad (35)$$

After some algebraic manipulation and gathering of geometric series terms we obtain the simplified expression:

$$\begin{aligned}
\check{P}_b^c(\tilde{\sigma}) \leq & \frac{1}{K} \sum_{d=2}^{\infty} \frac{\mathbf{A}^o(d) \overline{\mathbf{A}}^i(d)}{(2N)^d} \sum_{\mu=0}^1 N^{-\mu} e^{-\frac{(1-\alpha)^2 d - \mu \alpha^2}{\tilde{\sigma}^2}} \\
& 2N(d!) e^{-\frac{2\alpha}{\tilde{\sigma}^2}} {}_1F_1\left(1-d; 2-\mu; -2N e^{-\frac{2\alpha}{\tilde{\sigma}^2}}\right) \\
& - \frac{1}{K} \sum_{d=m_c+1}^{\infty} \frac{\mathbf{A}^o(d) \overline{\mathbf{A}}^i(d)}{(2N)^d} \sum_{\mu=0}^1 N^{-\mu} e^{-\frac{(1-\alpha)^2 d - \mu \alpha^2}{\tilde{\sigma}^2}} \\
& \frac{d!}{(m_c+1-\mu)!} \binom{d-1}{m_c} (2N e^{-\frac{2\alpha}{\tilde{\sigma}^2}})^{m_c+1} \\
& {}_2F_2\left(1, m_c+1-d; m_c+1, m_c+2-\mu; -2N e^{-\frac{2\alpha}{\tilde{\sigma}^2}}\right)
\end{aligned} \tag{36}$$

where ${}_2F_2$ is the generalized hypergeometric function [31].

3) *precoded TE*: By examining the precoded trellis in Fig. 3, the squared Euclidean distance of a compound error of Hamming distance d and length L for the precoded channel $\frac{1-\alpha D}{1 \oplus D}$ is:

$$d_E^2 = \lceil \frac{d}{2} \rceil + \lfloor \frac{d}{2} \rfloor \alpha^2 + 4\alpha\gamma + (1-\alpha)^2(L-d).$$

Substituting this expression in the bound on BER of the precoded TE derived in [11], and utilizing the approximation of the Q function once again, we obtain:

$$\begin{aligned}
\check{P}_b(\tilde{\sigma}) \leq & \frac{1}{K} \sum_{d=2}^N \frac{\mathbf{A}^o(d) \overline{\mathbf{A}}^i(d)}{\binom{N}{d}} \\
& \sum_{L=d}^N \left(\frac{1}{2}\right)^{L-d} \binom{N-L+\lfloor \frac{d}{2} \rfloor}{\lfloor \frac{d}{2} \rfloor} \binom{L-1-\lceil \frac{d-1}{2} \rceil}{\lfloor \frac{d-1}{2} \rfloor} \\
& e^{-\frac{(1-\alpha)^2(L-d) + \lceil \frac{d}{2} \rceil + \lfloor \frac{d}{2} \rfloor \alpha^2}{\tilde{\sigma}^2}} \sum_{\gamma=0}^{L-d} \binom{L-d}{\gamma} e^{-\frac{4\alpha\gamma}{\tilde{\sigma}^2}}.
\end{aligned} \tag{37}$$

Evaluating the summation over γ by utilizing the binomial identity, we obtain after some simplification:

$$\begin{aligned}\check{P}_b(\tilde{\sigma}) &\leq \frac{1}{K} \sum_{d=2}^N \frac{\mathbf{A}^o(d) \overline{\mathbf{A}}^i(d)}{\binom{N}{d}} e^{-\frac{[\frac{d}{2}] + \lfloor \frac{d}{2} \rfloor \alpha^2}{\tilde{\sigma}^2}} \\ &\quad \sum_{L=d}^N \binom{N-L+\lfloor \frac{d}{2} \rfloor}{\lfloor \frac{d}{2} \rfloor} \binom{L-1-\lceil \frac{d-1}{2} \rceil}{\lfloor \frac{d-1}{2} \rfloor} \Psi(\alpha, \tilde{\sigma})^{L-d} \\ \Psi(\alpha, \tilde{\sigma}) &= e^{-\frac{1+\alpha^2}{\tilde{\sigma}^2}} \cosh\left(\frac{2\alpha}{\tilde{\sigma}^2}\right)\end{aligned}\tag{38}$$

This can be simplified to a single sum over d by the utility of the generalized hypergeometric representation, which is given by:

$$\begin{aligned}\check{P}_b(\tilde{\sigma}) &\leq \frac{1}{K} \sum_{d=2}^N \frac{\mathbf{A}^o(d) \overline{\mathbf{A}}^i(d)}{\binom{N}{d}} \binom{N-\lceil \frac{d}{2} \rceil}{\lfloor \frac{d}{2} \rfloor} e^{-\frac{[\frac{d}{2}] + \lfloor \frac{d}{2} \rfloor \alpha^2}{\tilde{\sigma}^2}} \\ &\quad {}_3F_2\left(1, \lfloor \frac{d+1}{2} \rfloor, d-N; \lceil \frac{d}{2} \rceil - N, 1; \Psi(\alpha, \tilde{\sigma})\right).\end{aligned}\tag{39}$$

When $N \gg d$, we can use similar approximations to the ones used in the derivation of the interleaver gain exponent, by which one reaches a looser, albeit simpler, bound:

$$\begin{aligned}\check{P}_b(\tilde{\sigma}) &\leq \frac{1}{K} \sum_{d=2}^N \mathbf{A}^o(d) \overline{\mathbf{A}}^i(d) \frac{d!}{\lfloor \frac{d}{2} \rfloor!} N^{-\lceil \frac{d}{2} \rceil} e^{-\frac{[\frac{d}{2}] + \lfloor \frac{d}{2} \rfloor \alpha^2}{\tilde{\sigma}^2}} \\ &\quad {}_3F_2\left(1, \lfloor \frac{d+1}{2} \rfloor, d-N; \lceil \frac{d}{2} \rceil - N, 1; \Psi(\alpha, \tilde{\sigma})\right).\end{aligned}\tag{40}$$

REFERENCES

- [1] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. of IEEE ICC*, Geneva, Switzerland, May 1993, pp. 1740–1745.
- [2] S. Benedetto and G. Montorsi, "Unveiling turbo codes: Some results on parallel concatenated coding schemes," *IEEE Trans. on Inform. Theory*, vol. 42, pp. 409–428, March 1996.
- [3] C. Douillard, M. Jzquel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: Turbo-equalization," *Eur. Trans. Telecommun.*, vol. 6, pp. 507–511, Sep./Oct. 1995.
- [4] R. Koetter, A. Singer, and M. Tuchler, "Turbo equalization," *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 67–80, January 2004.

- [5] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. on Inform. Theory*, vol. 42, no. 2, pp. 429–445, March 1996.
- [6] W. Ryan, "Performance of high rate turbo codes on a PR4-equalized magnetic recording channel," *Proc. of IEEE ICC*, vol. 2, pp. 947–951, June 1998.
- [7] M. Reed, C. Schlegel, P. Alexander, and J. Asenstorfer, "Iterative multiuser detection for CDMA with FEC: Near-single-user performance," *IEEE Trans. on Communi.*, vol. 46, no. 12, pp. 1693–1699, December 1998.
- [8] T. V. Souvignier, M. Oberg, P. H. Siegel, R. E. Swanson, and J. K. Wolf, "Turbo decoding for partial response channels," *IEEE Trans. on Communi.*, vol. 48, no. 8, pp. 1297–1308, August 2000.
- [9] L. McPheters, S. McLaughlin, and K. Narayanan, "Precoded PRML, serial concatenation, and iterative (turbo) decoding for digital magnetic recording," *IEEE Trans. on Magn.*, vol. 35, no. 5, pp. 2325–2327, September 1999.
- [10] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding," *IEEE Trans. on Inform. Theory*, vol. 44, no. 3, pp. 909–926, May 1998.
- [11] M. Oberg and P. H. Siegel, "Performance analysis of turbo-equalized partial response channels," *IEEE Trans. on Communi.*, vol. 49, no. 3, pp. 436–444, March 2001.
- [12] J. Moon and J. Park, "Detection of prescribed error events: Application to perpendicular recording," *Proc. of IEEE ICC*, vol. 3, pp. 2057–2062, May 2005.
- [13] J. Park and J. Moon, "High-rate error-correction codes targeting dominant error patterns," *IEEE Trans. on Magn.*, vol. 42, no. 10, pp. 2573–2575, October 2006.
- [14] —, "A new class of error-pattern-correcting codes capable of handling multiple error occurrences," *IEEE Trans. on Magn.*, vol. 43, no. 6, pp. 2268–2270, June 2007.
- [15] —, "Error-pattern-correcting cyclic codes tailored to a prescribed set of error cluster patterns," *IEEE Trans. on Inform. Theory*, vol. 55, no. 4, pp. 1747–1765, April 2009.
- [16] H. Alhussien, J. Park, and J. Moon, "Iterative decoding based on error pattern correction," *IEEE Trans. on Magn.*, vol. 44, no. 1, pp. 181–186, January 2008.
- [17] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (corresp.)," *IEEE Trans. on Inform. Theory*, vol. 20, no. 2, pp. 284–287, March 1974.
- [18] J. Hagenauer and P. Hoeher, "A viterbi algorithm with soft-decision outputs and its applications," in *Proc. of IEEE GLOBECOM*, vol. 3, November 1989, pp. 1680–1686.
- [19] M. Tuchler, R. Koetter, and A. Singer, "Turbo equalization: principles and new results," *IEEE Trans. on Communi.*, vol. 50, no. 5, pp. 754–767, May 2002.
- [20] M. Tuchler, A. Singer, and R. Koetter, "Minimum mean squared error equalization using a priori information," *IEEE Trans. on Signal Processing*, vol. 50, no. 3, pp. 673–683, March 2002.
- [21] M. Oberg and P. H. Siegel, "Performance bound for parity-check coded partial-response channels," in *Proc. of IEEE ICC*, vol. 9, 2001, pp. 2701–2705.
- [22] H. Alhussien, "Channel matched iterative decoding for magnetic recording systems," Ph.D. dissertation, Univ. of Minnesota, Twin Cities, Minneapolis, April 2009.
- [23] R. Pyndiah, "Near-optimum decoding of product codes: Block turbo codes," *IEEE Trans. on Communi.*, vol. 46, no. 8, pp. 1003–1010, August 1998.
- [24] D. Chase, "Class of algorithms for decoding block codes with channel measurement information," *IEEE Trans. on Inform. Theory*, vol. 18, no. 1, pp. 170–182, January 1972.

- [25] S.-W. Lee and B. V. K. Vijaya Kumar, "Pattern-flipping chase-type decoders with error pattern extracting viterbi algorithm over partial response channels," *IEEE Journal on Sel. Areas in Communi.*, vol. 28, no. 2, pp. 181–187, February 2010.
- [26] H. Alhussien and J. Moon, "An iteratively decodable tensor product code with application to data storage," *IEEE Journal on Sel. Areas in Communi.*, vol. 28, no. 2, pp. 228–240, February 2010.
- [27] D. N. Rowitch, "Convolutional and turbo coded multicarrier direct sequence CDMA, and applications of turbo codes to hybrid ARQ communication systems," Ph.D. dissertation, Univ. California, San Diego, La Jolla, June 1998.
- [28] M. Chiani, D. Dardari, and M. Simon, "New exponential bounds and approximations for the computation of error probability in fading channels," *IEEE Trans. on Wireless Communi.*, vol. 2, no. 4, pp. 840–845, July 2003.
- [29] M. Petkovšek, H. S. Wilf, and D. Zeilberger, $A = B$. A K Peters, Ltd, 1997.
- [30] L. J. Slater, *Confluent hypergeometric functions*. Cambridge University Press, 1960.
- [31] —, *Generalized hypergeometric functions*. Cambridge University Press, 1966.